

Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN  
INGEGNERIA STRUTTURALE E IDRAULICA

Ciclo XXIV

Settore Concorsuale di afferenza: 08/A1

Settore Scientifico disciplinare: ICAR/02

TITOLO TESI

VARIABILITA' SPAZIALE DEI PARAMETRI DI  
MODELLI AFFLUSSI-DEFLUSSI

**Elena Montosi**

**Coordinatore Dottorato:**

**Relatore:**

**Chiar.mo Prof. Ing. Erasmo Viola**

**Prof. Ing. Alberto Montanari**

**Esame finale anno 2012**



# Indice:

<b>Premessa</b> .....	1
-----------------------	---

## **PARTE I:**

### **Analisi della Variabilità Spaziale dei Parametri di Modelli Afflussi-Deflussi e Prima Verifica dell'Ipotesi di omogeneità Parametrica**

1	Il modello AFFDEF .....	5
1.1	Dati in ingresso: .....	5
1.2	Accumulo e scioglimento nivale .....	6
1.2.1	Background.....	6
1.2.2	Il modulo neve .....	9
1.3	Intercettazione .....	10
1.4	Evapotraspirazione .....	11
1.5	Infiltrazione .....	12
1.6	Propagazione dei deflussi superficiali e profondi .....	15
1.7	I parametri del modello AFFDEF .....	17
2	L'incertezza nei modelli.....	19
2.1	La calibrazione dei modelli afflussi-deflussi.....	20
2.2	Equifinalità .....	23
2.3	Stima dell'incertezza parametrica: .....	25
2.3.1	Il teorema di Bayes: .....	25
2.3.2	L'algoritmo Random Walk Metropolis .....	26
2.4	L'algoritmo Differential Evolution Adaptive Metropolis (DREAM) .....	28
3	Caso di studio: Stima dei parametri del modello AFFDEF .....	32
3.1	Area di studio .....	32
3.2	Scelta del modello: .....	33
3.3	Informazione disponibile.....	34
3.3.1	Informazioni idro-climatiche .....	34
3.4	Stima dell'incertezza .....	35
3.4.1	Analisi dei risultati delle simulazioni: il caso del torrente Leo .....	37
3.5	Discussione dei risultati .....	42
3.6	Analisi dell'ipotesi di invarianza spaziale dei parametri.....	47
4	Stima dell'approssimazione indotta dall'adozione di parametri omogenei nello spazio. 49	
4.1	Bacini non strumentati .....	50
4.2	L'ipotesi di omogeneità dei parametri: la procedura di calibrazione .....	51
4.3	Discussione dei risultati .....	54
5	Considerazioni conclusive della I parte.....	59

## **PARTE II:**

### **Analisi Regionale dell'Ipotesi di omogeneità Parametrica**

Introduzione .....	61
1 Cross-calibrazione di modelli afflussi-deflussi a scala regionale.....	63
2 Modelli idrologici e aree oggetto di studio .....	65
2.1 La regione dell' Innviertel.....	65
2.2 Il contesto italiano: i bacini dei fiumi Enza, Secchia, Panaro e Reno .....	67
2.3 Scelta del modello afflussi-deflussi.....	68
2.4 Il modello HBV.....	70
2.5 La versione del modello HBV adottata nella presente trattazione .....	74
2.6 Il modello HYMOD .....	77
2.7 Misura di similitudine idrologica: distanza di Ghosh.....	78
3 Problematiche legate alla scelta del periodo di calibrazione.....	80
3.1 Lunghezza del periodo di calibrazione.....	82
3.2 Non contemporaneità delle osservazioni di portata su bacini differenti .....	87
4 Cross-calibrazione ed incertezza nei dati .....	90
4.1 Modelli idrologici e incertezza nei dati.....	90
4.2 Cross-calibrazione su bacini non annidati.....	95
5 Considerazioni conclusive della II parte .....	100
<b>Appendice .....</b>	<b>103</b>
<b>Bibliografia .....</b>	<b>110</b>
<b>Elenco delle Figure .....</b>	<b>115</b>
<b>Elenco delle Tabelle.....</b>	<b>118</b>

## Premessa:

L'idrologia è stata definita da Penman (1961) come la scienza che cerca di rispondere alla domanda "che cosa succede quando piove?"; i modelli idrologici sono dunque i modelli matematici che vengono ideati per rispondere a questa domanda con un grado di dettaglio che dipende dal problema pratico con il quale ci si vuole confrontare (Singh, 1995). Applicazioni che riguardano, ad esempio, la previsione delle piene o la stima di disponibilità di risorsa idrica fanno ricorso a simulazioni idrologiche che possono essere un valido supporto anche per studi ecologici. Lo studio presentato in questo lavoro di tesi vuole porre l'attenzione anzitutto sul problema della stima del bilancio idrico a scala regionale; a tutt'oggi infatti sono numerose le aree del globo nelle quali i corsi d'acqua sono dotati di scarse, inadeguate o poco efficienti strumentazioni per la valutazione della portata (Sivapalan et al, 2003) ed in tali condizioni, non conoscendo l'ammontare di risorsa a disposizione, risulta difficoltoso, ad esempio, effettuarne una corretta ripartizione tra tutti gli usi. La simulazione idrologica può essere dunque un valido dispositivo per ottenere una stima del bilancio idrico a livello regionale.

Esistono numerosi tipi di modelli afflussi-deflussi che si differenziano per struttura e tipo di applicazione alla quale sono destinati; in quanto a complessità poi si spazia dai parsimoniosi modelli concettuali a parametri concentrati alle complesse rappresentazioni realizzate con modelli fisicamente basati e spazialmente distribuiti (Yadav et al., 2007). Qualunque modello si adotti per le applicazioni pratiche, è sempre necessaria la preliminare calibrazione dei suoi parametri (Beven, 2001), procedura con la quale si sceglie il valore da attribuire ai parametri del modello, così che la risposta del bacino simulata dal modello e la serie di portate osservate siano il più simili possibile. Anche i modelli fisicamente basati necessitano, in qualche misura, della calibrazione di alcuni parametri, poiché risulta difficile stimare il valore di tutti i parametri ricorrendo a misure effettuate sul campo (Beven, 2001); la scala alla quale vengono realizzate le misurazioni è infatti molto inferiore rispetto a quella alla quale si applicano i modelli afflussi deflussi e le simulazioni idrologiche sono solo il risultato della generalizzazione dei processi che avvengono a piccola scala (Bergstrom, 1991). La fase di calibrazione dei parametri di un modello afflussi-deflussi ha dunque un ruolo importante per la generazione di simulazioni idrologiche di buona qualità.

Si può scegliere di ricorrere all'uso di un modello di tipo spazialmente distribuito per rappresentare in modo esplicito la variabilità dei fenomeni che concorrono alla generazione di deflusso in un bacino; questi modelli prevedono la suddivisione dell'area di studio in unità fondamentali, di dimensioni ridotte in modo tale che l'ipotesi di omogeneità dei processi idrologici che avvengono al suo interno sia ragionevolmente plausibile. Le equazioni costitutive del modello vengono applicate al volume di controllo che ricade all'interno di ciascuna unità. Da un lato questi modelli portano al grande vantaggio che, una volta calibrati nella sezione di chiusura del bacino, è possibile produrre simulazioni idrologiche in qualunque sezione fluviale, anche non strumentata, posta al suo interno senza la necessità di ripetere la procedura di calibrazione (Moretti & Montanari, 2006); d'altro canto però, proprio il frazionamento dell'area

di studio in numerose entità elementari comporta un importante sforzo computazionale per la necessità di applicare le equazioni fondamentali su ciascuna unità.

L'obiettivo che ci si è posti è la simulazione del bilancio idrico a scala regionale, dunque si prevede la necessità di produrre simulazioni idrologiche simultaneamente in diverse sezioni della rete drenante, non collocate all'interno del medesimo bacino idrologico. E' doveroso perciò ripetere l'operazione di calibrazione del modello su ciascun bacino, procedura che richiede tempi di calcolo consistenti. Per cercare di ridurli si può ricorrere alla diminuzione del numero di parametri coinvolti nella procedura di calibrazione stessa, ipotizzando che, per alcuni parametri, sia plausibile l'ipotesi di omogeneità su tutta la regione studiata. Per avallare questa ipotesi si procede allo studio della variabilità spaziale dei parametri stessi ricorrendo ad analisi di incertezza parametrica.

Un metodo per ottenere simulazioni idrologiche in corrispondenza di sezioni scarsamente strumentate alternativo all'utilizzo di modelli idrologici spazialmente distribuiti è il ricorso ad un'analisi di tipo regionale. In questo caso si è soliti adottare modelli concettuali a parametri concentrati. L'utilizzo di tali modelli presuppone la consapevolezza che i processi fisici responsabili della formazione di deflusso hanno una variabilità estremamente elevata sull'area racchiusa all'interno del bacino stesso la cui conoscenza può essere solo limitata; dunque è possibile giustificare il ricorso a modelli dalla struttura semplice che vogliono dare alla rappresentazione di tali fenomeni un approccio sostanzialmente di tipo statistico (*Bergstrom, 1991*). Gli approcci regionali solitamente sviluppati in letteratura prevedono che un modello afflussi-deflussi venga calibrato in corrispondenza di un sostanzioso numero di sezioni fluviali strumentate. Si sviluppano poi equazioni matematiche con le quali si cerca di relazionare ciascun parametro del modello ad attributi geo-morfo-climatici caratteristici sia del bacino non strumentato che delle aree nella quali si dispone di misure di portata. Si tratta ancora una volta di processi lunghi e dispendiosi in termini di tempi di calcolo.

Nella seconda parte di questo lavoro di tesi si sviluppa dunque un'analisi tesa a valutare l'applicabilità di un approccio regionale differente da quelli sin ad ora presentati in letteratura; si vuole infatti proporre una calibrazione regionale nella quale non si ricorra all'uso delle caratteristiche proprie dei bacini per l'individuazione dei singoli parametri del modello. Si suppone infatti che l'utilizzo di informazioni provenienti da diversi bacini idrologici, a tutti gli effetti simili da un punto di vista idrologico al bacino in esame, porti all'individuazione di un insieme di parametri robusto. La procedura proposta prevede poi il trasferimento diretto al bacino non strumentato del set di parametri individuato con la calibrazione regionale, senza la necessità di ricorrere a regressioni regionali. Nella sostanza dunque si assume che l'intero set di parametri possa essere considerato costante su un'area costituita sia dai bacini strumentati che si adottano per la calibrazione del modello e sul bacino non strumentato al quale vengono direttamente trasferiti i parametri.

Il lavoro di tesi viene dunque strutturato in due differenti sezioni. La prima dedicata allo studio della variabilità spaziale dei parametri di un modello afflussi-deflussi spazialmente distribuito che viene sviluppata ricorrendo ad un'analisi di incertezza parametrica. Al termine della prima parte viene condotto un primo esperimento di verifica dell'impatto, in termini di approssimazione dei risultati, dell'ipotesi di invarianza spaziale dei parametri. Poiché i risultati

ottenuti sembrano meritevoli di ulteriori approfondimenti, nella seconda parte ci si concentra sull'analisi di tipo regionale, la quale è fondata su detta ipotesi di invarianza spaziale e in riferimento alla quale si è condotta una verifica di dettaglio del livello di approssimazione introdotto.





PARTE I:  
ANALISI DELLA VARIABILITÀ SPAZIALE DEI  
PARAMETRI DI MODELLI AFFLUSSI-DEFLUSSI E  
PRIMA VERIFICA DELL'IPOTESI DI OMOGENEITÀ  
PARAMETRICA

---



# 1 Il modello AFFDEF

La volontà di stimare, istante per istante, la portata defluente in ogni punto della rete drenante può essere soddisfatta ricorrendo a modelli afflussi-deflussi spazialmente distribuiti che, per loro natura, prevedono l'applicazione delle equazioni costitutive a volumi di controllo di dimensioni particolarmente ridotte; in tal modo è plausibile assumere omogenei i processi idrologici che si verificano al loro interno ed è possibile descrivere dettagliatamente la variabilità spaziale dei fenomeni che influenzano la formazione dei deflussi superficiali. Volendo analizzare un'ampia area di studio, è ragionevole propendere per l'applicazione di schemi di tipo concettuale piuttosto che per un approccio strettamente fisicamente basato. Quest'ultimo infatti richiederebbe una vasta mole di dati, di alta qualità, dotati di una risoluzione spaziale e temporale molto fine per modellare in modo attendibile i processi che avvengono nell'area (*Beven, 1989*). La rappresentazione concettuale comporta invece una schematizzazione più semplificata e meno dettagliata dei singoli fenomeni fisici in gioco, non andando però ad agire sulla rappresentazione della variabilità spaziale di tutte le caratteristiche geo-morfo-climatiche proprie dell'area analizzata.

Il modello idrologico scelto per la prima parte dello studio è dunque un modello a simulazione continua, spazialmente distribuito che fa uso di una schematizzazione in larga parte concettuale dei processi idrologici. Sviluppato da *Moretti & Montanari (2007)*, il modello, denominato AFFDEF, è stato originariamente pensato per essere applicato a scala di bacino. Poiché però si vuole stimare il bilancio idrico di una vasta area, sono state indispensabili alcune modifiche; in particolare si è modificata leggermente la struttura, passando da un intervallo di calcolo orario a quello giornaliero, ripensando il metodo di propagazione dei deflussi superficiali e profondi ed inserendo inoltre un modulo neve in modo da rappresentare al meglio ciò che accade nei bacini montani.

Il modello AFFDEF discretizza l'area di studio in celle coincidenti con i pixels del modello digitale del terreno, per ognuna delle quali si stima il bilancio idrologico locale. All'interno di ogni cella si considera infatti l'interazione tra suolo, vegetazione ed atmosfera simulando i fenomeni di accumulo e scioglimento nivale, intercettazione di parte della pioggia, evapotraspirazione, infiltrazione e ruscellamento.

## 1.1 Dati in ingresso:

Le altezze di precipitazione (mm) e le temperature medie atmosferiche (°C) registrate in diverse stazioni pluvio-termometriche presenti nell'area di interesse sono i dati che devono essere forniti in ingresso al modello con il medesimo passo temporale (giornaliero nel caso preso in esame); maggiore è il numero di stazioni di misura distribuite omogeneamente sul territorio e migliore risulta la rappresentazione della variabilità spaziale dei fenomeni. In

AFFDEF è implementato un modulo che stima la precipitazione  $P[t, (i, j)]$  (mm), da considerare istante per istante su ogni cella di coordinate  $(i, j)$  attraverso due differenti metodi: è possibile scegliere alternativamente tra il metodo dei poligoni di Thiessen (*Thiessen, 1911*) e il metodo di interpolazione della inversa distanza al quadrato (*Moretti & Montanari, 2007*). Mentre con la prima metodologia si associa ad ogni cella del bacino, nella quale non sia stata registrata alcuna misura di precipitazione, il valore fornito dal pluviometro ad essa più prossimo, con la seconda si procede ad una interpolazione delle misure fornite dai diversi strumenti di misura presenti nell'area. La stima della temperatura media giornaliera locale si rifà unicamente al metodo dei poligoni di Thiessen e prevede anche il ragguaglio della misura stessa alla quota effettiva della cella  $(i, j)$ .

Le informazioni di altimetria della zona di studio sono fornite ad AFFDEF dal modello digitale del terreno, mentre la caratterizzazione spaziale della capacità di infiltrazione del suolo dalla mappa del Curve Number (*Soil Conservation Service, 1972*).

## **1.2 Accumulo e scioglimento nivale**

### **1.2.1 Background**

Radiazione solare netta (radiazione solare a onda lunga e a onda corta), trasferimento di calore sensibile e latente e calore associato alla precipitazione sono le principali componenti dello scambio di energia che avviene tra aria e manto nevoso. La quantità di radiazione solare incidente varia al variare sia della stagione che del momento della giornata, è influenzato dalla presenza o meno di copertura nuvolosa e dal tipo di vegetazione, ed è differente a seconda di pendenza ed aspect del versante oggetto di studio (*Pellicciotti et al., 2005*). La neve è caratterizzata da albedo alta, cioè riflette verso l'atmosfera la maggior parte della radiazione solare ad onda corta incidente. L'albedo tipica della neve fresca può infatti raggiungere il 90%, ma all'aumentare del grado di invecchiamento dello strato di neve si può avere un calo di albedo, indice quindi di un maggiore assorbimento di radiazione. La quantità di radiazione ad onda lunga emessa da un oggetto che si trova in atmosfera o sulla superficie della terra (ad es. alberi) è legata sia all'emissività dell'oggetto stesso, sia alla temperatura della sua superficie. Anche il manto nevoso è caratterizzato da alti valori di emissività. I trasferimenti di calore latente e sensibile sono processi altamente turbolenti e, in quanto tali, la loro entità risulta altamente legata alla velocità del vento che spira. Quando poi si verifica un evento di pioggia il trasferimento di calore avviene non più solo tra aria e neve, ma anche tra pioggia e neve. Il calore proprio della pioggia viene ceduto alla neve fino a quando si raggiunge un equilibrio con la temperatura dello strato nevoso stesso. La quantità di calore trasferito è funzione quindi sia della temperatura della pioggia, sia della quantità di pioggia stessa e provoca lo scioglimento di una frazione dello strato di neve. Si hanno quindi condizioni favorevoli allo scioglimento quando lo scambio di calore tra aria e neve è positivo e quando la superficie dello strato di neve si trova a 0°C. Solitamente ciò si verifica quando la temperatura atmosferica è positiva oppure anche quando, pur avendo temperature negative, ci si trova in presenza di particolari condizioni meteorologiche (*Anderson, 2006*).

In letteratura sono presenti due famiglie di metodi per la modellazione numerica dei processi di accumulo e scioglimento nivale:

- i modelli che si basano sulla stima del bilancio energetico,
- i modelli noti come “gradi-giorno” che assumono esista una relazione empirica tra la temperatura dell’aria e la velocità di scioglimento del manto nevoso.

Il modulo di scioglimento nivale inserito all’interno del modello ARNO (*Todini, 1996*), ad esempio, si basa sul paragone tra la stima dell’energia totale disponibile al passo temporale  $t$  e l’energia necessaria per lo scioglimento di almeno una parte della quantità di neve immagazzinata nel volume di controllo allo stesso istante temporale. La radiazione solare incidente, la cui entità viene stimata in base a misure di temperatura dell’aria e che, per considerare anche il ruolo giocato dall’albedo, viene rappresentata in modo differente a seconda della presenza o meno di nubi in cielo, è la prima delle componenti considerate nel calcolo dell’energia totale disponibile. Si considera poi anche la quota di energia associata alla precipitazione caduta in forma solida o liquida. Solo se l’energia totale disponibile è superiore all’energia effettivamente necessaria per mantenere la quantità di acqua immagazzinata allo stato solido alla temperatura di  $0^{\circ}\text{C}$ , si ha scioglimento di parte dell’ammasso nevoso.

Si sceglie invece di ricorrere ad un modello del tipo “gradi-giorno” soprattutto quando c’è una ridotta disponibilità di dati. Hock (*2003*) presenta una sintesi dei molteplici casi di applicazione di questo tipo di modelli, includendo anche lo studio di ghiacciai. Sottolinea che le due fonti principali di energia responsabili dello scioglimento nivale (radiazione infrarossa e flusso di calore sensibile) sono altamente influenzate dalla temperatura atmosferica, dando dunque una spiegazione fisica del successo della temperatura come indicatore unico dell’energia necessaria per lo scioglimento del manto nevoso.

Il più semplice dei modelli “gradi-giorno” lega lo scioglimento nivale,  $M$  (mm), che avviene in un periodo pari a  $n$  intervalli  $\Delta t$ , alla somma delle temperature atmosferiche positive  $T^+$  ( $^{\circ}\text{C}$ ) che si registrano in ogni intervallo  $\Delta t$ , attraverso un fattore di proporzionalità denominato *degree-day factor*, DDF, espresso in  $\text{mm}/(d \cdot ^{\circ}\text{C})$ .

$$\sum_{i=1}^n M = DDF \sum_{i=1}^n T^+ \Delta t \quad (1.1)$$

L’intervallo  $\Delta t$  è solitamente considerato pari a un giorno, anche se in linea di principio si può considerare qualsiasi altra discretizzazione temporale (passo orario o addirittura mensile).

E’ bene però notare che l’uso di dati di temperatura media giornaliera può essere fuorviante soprattutto quando si considerano i periodi in cui questa fluttua attorno allo zero; temperature medie giornaliere minori di zero indicherebbero quindi assenza di scioglimento, mentre nella realtà in gran parte della giornata possono essere prevalse condizioni che hanno portato allo scioglimento di una porzione di manto nevoso. E’ quindi auspicabile, secondo Hock (*2003*), l’adozione di un’alta risoluzione temporale per evitare la sovrastima del *degree-day factor*.

Nella pratica spesso la formulazione degree-day assume la forma:

$$M = \begin{cases} f_m(T_d - T_0), & T_d > T_0 \\ 0, & T_d \leq T_0 \end{cases} \quad (1.2)$$

nella quale  $M$  rappresenta lo scioglimento giornaliero,  $T_d$  è la temperatura media giornaliera,  $T_o$  è la temperatura di soglia al di sopra della quale si innesca lo scioglimento e  $f_m$  è il fattore di scioglimento. L'equazione (1.2) si differenzia dall'equazione (1.1) per l'utilizzo di una temperatura di soglia  $T_o$  che rende conto del fatto che nella realtà lo scioglimento di neve può avvenire non necessariamente per temperature superiori a  $0^\circ\text{C}$ .

Nel corso degli anni si sono susseguiti tentativi per apportare migliorie a questa formulazione, soprattutto giustificate dalla volontà di considerare anche l'effetto delle variabili vento, pressione di vapore e radiazione netta sul fenomeno dello scioglimento. E' stata questa la molla che in alcuni casi (*Anderson, 2006; Pellicciotti et al., 2005*) ha spinto verso approcci di graduale transizione tra espressioni di bilancio energetico e modelli gradi-giorno semplici, aumentando di conseguenza anche il numero di variabili in ingresso al modulo neve. Pellicciotti et al. (2005) per esempio inglobano all'interno della formulazione temperatura-indice, utilizzata a passo orario, anche la rappresentazione dell'albedo  $\alpha$ , calcolata come funzione logaritmica di un particolare tipo di temperatura, e della radiazione a onda corta  $G$  ( $\text{W m}^{-2}$ ), modificando così la classica formulazione dello scioglimento:

$$M = \begin{cases} TF \cdot T + SRT(1 - \alpha)G, & T > T_T \\ 0, & T \leq T_T \end{cases} \quad (1.3)$$

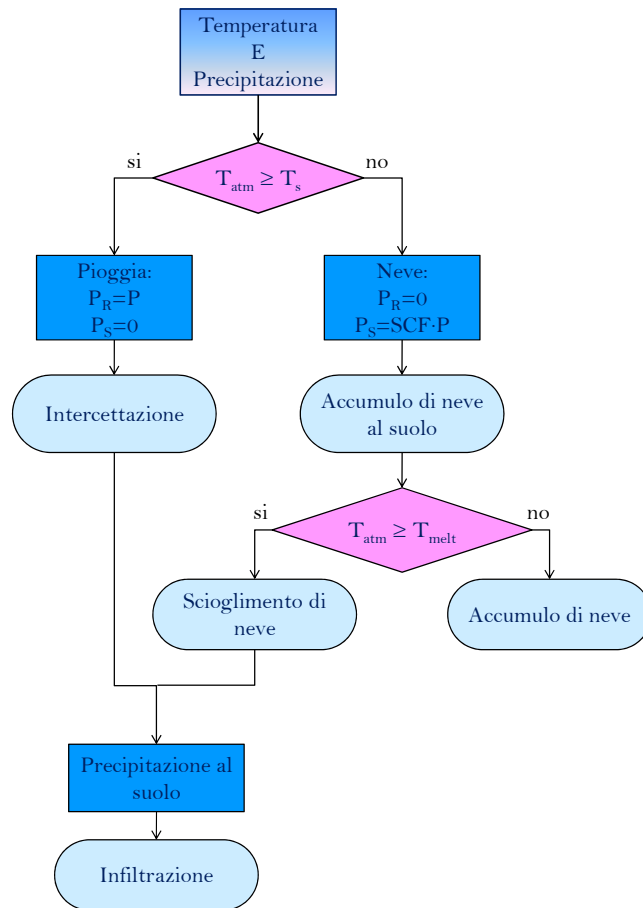
dove  $TF$  e  $SRT$  sono due coefficienti empirici espressi rispettivamente in ( $\text{mm h}^{-1} ^\circ\text{C}^{-1}$ ) e in ( $\text{m}^2 \text{m W}^{-1} \text{h}^{-1}$ ) e la temperatura  $T_T$  al di sotto della quale non avviene scioglimento è assunta pari a  $-1^\circ\text{C}$ . Questa applicazione prevede dunque la netta separazione di due contributi: nel primo termine, dipendente direttamente dalla temperatura atmosferica ricadono gli effetti della radiazione ad onda lunga e dei processi turbolenti, mentre il secondo termine considera solo la radiazione ad onda corta.

Un esempio ancora più esplicito di come si possano fondere l'approccio energetico e quello gradi-giorno è fornito da *Anderson (2006)* che nel modello SNOW-17 sceglie di calcolare lo scioglimento durante i periodi di assenza di pioggia con una formulazione che si rifà alla formulazione gradi-giorno, mentre simula l'effetto della precipitazione liquida su neve ricorrendo ad un approccio di bilancio energetico.

Esistono poi anche altri modelli che si discostano dalla formulazione base gradi-giorno; taluni vogliono considerare la significativa dipendenza del fattore scioglimento dalle caratteristiche morfo-pedologiche dell'area di studio, sottolineandone la significativa variabilità a scala locale, altri invece sottolineano la forte variazione stagionale dei fenomeni in gioco. Tra i primi si ricorda il modello HBV che, secondo la formulazione di *Lindstrom et al. (1997)*, modella in modo differente le aree boscate da quelle aperte e che prevede la suddivisione del bacino in zone di altitudine differente su ciascuna delle quali calibra i parametri del modello. *Anderson (2006)* decide invece di considerare il fattore scioglimento variabile a seconda del periodo dell'anno. Esistono poi numerose altre formulazioni; talune ancora più semplici rispetto al modello gradi-giorno, come quella adottata da *Arnold et al. (1998)*, altre invece che prevedono l'utilizzo combinato di dati di copertura nevosa derivati da satellite e di misurazioni al suolo (*Rango, 1995; Parajka & Blöschl, 2008*).

## 1.2.2 Il modulo neve

Consapevoli della complessità dei fenomeni fisici che sono coinvolti nei processi di accumulo e scioglimento della neve, si sceglie di adottare un modello che nella sua semplicità porti ad una buona rappresentazione della realtà fisica. Si adotta perciò una formulazione gradi-giorno, molto simile a quella introdotta inizialmente da Lindstrom et al. (1997) nel modello HBV e poi modificata, tra gli altri, da Parajka et al. (2007). La necessità di mantenere una formulazione il più semplice possibile e la consapevolezza che il modello verrà applicato allo studio di bacini situati nella zona appenninica della regione Emilia Romagna, caratterizzati quindi dalla presenza di neve solitamente solo ad alta quota, ha portato alla scelta di adottare gli stessi parametri per tutta l'area di studio. Non si prevede quindi la necessità di frazionare i bacini in zone a differente quota, né a differente copertura vegetale. In Figura I- 1 si presenta il diagramma di flusso che sta alla base del modulo neve utilizzato.



**Figura I- 1:** Rappresentazione schematica della simulazione dei processi di accumulo e scioglimento nivale

La precipitazione totale giornaliera  $P$  (mm) è considerata come pioggia,  $P_R$  (mm), se la temperatura giornaliera media misurata,  $T_{atm}$  (°C), risulta superiore ad una temperatura di soglia  $T_s$  (°C). Per temperature rigide, la precipitazione è invece considerata completamente come neve, senza prevedere la possibilità di una fase di transizione di coesistenza di una frazione di pioggia e di una di neve.

$$\begin{aligned}
&\text{se } T_{atm} \geq T_S \text{ allora } \begin{cases} P_R = P \\ P_S = 0 \end{cases} \\
&\text{se } T_{atm} < T_S \text{ allora } \begin{cases} P_R = 0 \\ P_S = SCF \cdot P \end{cases}
\end{aligned} \tag{1.4}$$

Le precipitazioni nevose sono spesso accompagnate da vento e questo provoca un deficit nella misura effettiva realizzata dagli strumenti a terra (*Parajka et al., 2007*); a ciò si aggiunge il fatto che solitamente la rete di misura non dispone di un numero adeguato di nivometri o pluviometri riscaldati, ma che spesso le misure di precipitazione sono unicamente realizzate con pluviometri. Gli errori nella misura delle precipitazioni nevose sono dunque solitamente ingenti e si rende pertanto necessaria l'introduzione di un fattore moltiplicativo *SCF* adimensionale che li consideri.

Lo scioglimento è modellato secondo il classico approccio gradi-giorno ed avviene solo quando la temperatura atmosferica è superiore al valore assunto da  $T_{melt}$  ( $^{\circ}\text{C}$ ).

$$M_{snow} = M_f \cdot (T_{atm} - T_{melt}) \text{ se } T_{atm} > T_{melt} \text{ e se } SWE > 0 \tag{1.5}$$

nella quale  $M_{snow}$  [L] è la quantità di neve che si scioglie ad ogni passo temporale e  $M_f$  è il fattore gradi-giorno [L/T/t] solitamente espresso come (mm/ $^{\circ}\text{C}$ /d). La variazione tra l'equivalente in acqua dello strato di neve presente il giorno  $i$ ,  $SWE_i$  [L], e l'equivalente in acqua del giorno precedente,  $SWE_{i-1}$ , si calcola mediante la formula:

$$SWE_i = SWE_{i-1} + (P_S - M_{snow}) \cdot \Delta t \tag{1.6}$$

nella quale  $\Delta t$  è l'intervallo di simulazione pari a 1 giorno.

### 1.3 Intercettazione

Parte della precipitazione può essere intercettata e trattenuta dalla vegetazione, in quantità variabile a seconda del tipo di copertura fogliare e dell'intensità e durata della pioggia. Boschi, prati e pascoli portano a una differente entità di intercettazione poiché sono caratterizzati da composizione, età, densità e grado di copertura completamente differenti. Precipitazioni di breve durata e lieve entità possono portare anche a intercettazione totale, mentre nel caso di piogge intense e di lunga durata, la frazione trattenuta dalla vegetazione può tendere a zero. Contemporaneamente all'intercettazione si verifica anche l'evaporazione di parte del volume idrico bloccato sulle foglie. La porzione di pioggia non intercettata dalla copertura vegetale arriva invece al suolo direttamente o scorrendo lungo i rami e i fusti.

Il fenomeno dell'intercettazione viene descritto per ogni cella  $(i,j)$  in cui viene discretizzato il bacino oggetto di studio attraverso l'uso di un serbatoio locale a livello della vegetazione, definito serbatoio di intercettazione, nel quale viene trattenuta unicamente la prima frazione di precipitazione liquida. Questo fenomeno non viene infatti considerato nel caso di precipitazioni nevose. La capacità del serbatoio in questione è differente per ogni cella ed è pari a  $C_{int} \cdot S(i,j)$ , essendo  $C_{int}$  un parametro che necessita di essere calibrato ed  $S(i,j)$  la capacità di



immagazzinamento locale. Si rimanda al paragrafo 1.5 per una dettagliata descrizione del modello.

## 1.4 Evapotraspirazione

L'evaporazione consiste nella restituzione in atmosfera, sotto forma di vapore, dell'acqua presente nei corpi idrici superficiali, nel terreno e sulla superficie della copertura vegetale. Si parla invece di traspirazione per definire l'estrazione di acqua dal terreno per mezzo delle radici delle piante e il successivo rilascio in atmosfera attraverso le loro foglie. Si è soliti riferirsi alla combinazione dei due processi con il termine evapotraspirazione. Tale fenomeno, seppure di entità trascurabile per eventi piovosi brevi e di forte intensità, influenza significativamente il bilancio idrologico nel lungo periodo, provocando la significativa diminuzione di umidità all'interno del suolo ad una velocità determinata proprio dall'entità dell'evapotraspirato.

L'evapotraspirazione effettiva indica invece le perdite solitamente riscontrate in condizioni reali, dipendenti quindi dalle condizioni atmosferiche e dalle caratteristiche proprie della superficie evaporante (tipo, densità, grado di copertura vegetale, tipo di suolo e caratteristiche dei corpi idrici superficiali).

In condizioni di disponibilità idrica non limitante, l'evapotraspirazione da un terreno completamente ricoperto di vegetazione omogenea, in buono stato vegetativo, è determinata solo dalle condizioni meteorologiche ed in parte dal tipo di coltivazione; in queste condizioni standard l'evapotraspirazione prende il nome di evapotraspirazione potenziale. Il servizio Arpa SIMC della regione Emilia Romagna (*R.E.R., 2010*) ha stimato che l'evapotraspirazione potenziale annua in regione vari tra valori prossimi ai 1100 mm annui, caratteristici di alcune zone pianeggianti in provincia di Ravenna e Modena, e i 500 mm annui stimati per le aree del crinale appenninico che si trovano a quota maggiore.

L'algoritmo adottato in AFFDEF è noto come *radiation method* (Doorenbos & Pruitt, 1975), si basa sulla semplificazione dell'equazione di Penman-Monteith (Maidment, 1992) e, poiché gli effetti del fenomeno evapotraspirazione sono evidenti soprattutto sul lungo periodo, si giustifica il ricorso ad una trattazione che trascura i termini legati alla pressione di vapore e alla velocità del vento, non essendo questi solitamente di facile reperimento.

L'equazione empirica che stima l'evapotraspirazione potenziale  $ET_{pm}[t, (i, j)]$  calcolata su base mensile e riferita alla cella  $(i, j)$  è espressa come una funzione lineare della temperatura.

$$ET_{pm}[t, (i, j)] = \alpha + \beta \cdot N(k) \cdot W_{TA}(k) \cdot T_m[t, (i, j)] \quad (1.7)$$

nella quale:

- $\alpha$  (mm) e  $\beta$  ( $\text{mm}^\circ\text{C}^{-1}$ ) sono i coefficienti di regressione calcolati secondo il *radiation method*,
- $T_m[t, (i, j)]$  ( $^\circ\text{C}$ ) è la temperatura osservata al tempo  $t$ ,
- $N(k)$  (-) è la media mensile, riferita al mese  $k$ , del numero massimo di ore di sole giornaliere tabulato in funzione della latitudine,
- $W_{TA}(k)$  (-) è il fattore di compensazione.

L'equazione (1.7) può essere utilizzata anche a scala temporale inferiore per stimare l'evapotraspirazione potenziale  $ET_{pd}[t, (i, j)]$  [L/t] nel caso in questione a passo temporale giornaliero. Nel caso oggetto di studio la temperatura  $T_m$  è quindi la temperatura media giornaliera osservata ed è necessario dividere tutta l'equazione per  $n(k)$ , cioè il numero di giorni presenti nel mese in questione:

$$ET_{pd}[t, (i, j)] = \{\alpha + \beta \cdot N(k) \cdot W_{TA}(k) \cdot T_m[t, (i, j)]\} / n(k) \quad (1.8)$$

Per maggiori dettagli si rimanda il lettore ai contenuti dell'appendice A.

## 1.5 Infiltrazione

Quando la precipitazione liquida entra in contatto con la superficie del terreno vi penetra, almeno in parte, per forza gravitazionale e per capillarità. La velocità con la quale si verifica questo fenomeno dipende da fattori che rimangono costanti nel tempo, come la tessitura del suolo (terreni di tessitura fine sono caratterizzati da particelle di granulometria fine tra le quali si hanno piccoli pori e dunque bassa la velocità di infiltrazione), ma anche da fattori che possono variare nel tempo, primo tra tutte l'intensità della precipitazione stessa. Il contenuto di umidità già presente all'interno del suolo gioca ad esempio un ruolo fondamentale; il tasso di infiltrazione dell'acqua diminuisce all'aumentare del contenuto di umidità del suolo. Inoltre si ricorda che la pioggia si infila con maggiore facilità all'interno di suoli granulari piuttosto che in suoli compatti; poiché la struttura del terreno è altamente influenzata dalle pratiche di coltivazione adottate, queste condizionano anche l'infiltrazione propria del suolo stesso.

La conducibilità idraulica  $k$  [L/t] e la capacità di ritenzione di volume idrico per unità di volume di suolo  $C_s$  [-] sono le due caratteristiche idrauliche fondamentali attraverso le quali si rappresenta il fenomeno dell'infiltrazione.

In AFFDEF la separazione tra i deflussi superficiali e quelli profondi viene effettuata ricorrendo al metodo del *Curve Number* modificato che adotta una schematizzazione concettuale simile a quella del metodo SCS-CN (*Soil Conservation Service, 1972*). Moretti e Montanari (2007) scelgono di affidarsi alla praticità del metodo SCS-CN, che però era stato messo a punto per simulazioni idrologiche a scala di evento. Necessitando di un modello che produca simulazioni in continuo del contenuto idrico del terreno anche durante i periodi secchi, hanno introdotto alcune modifiche.

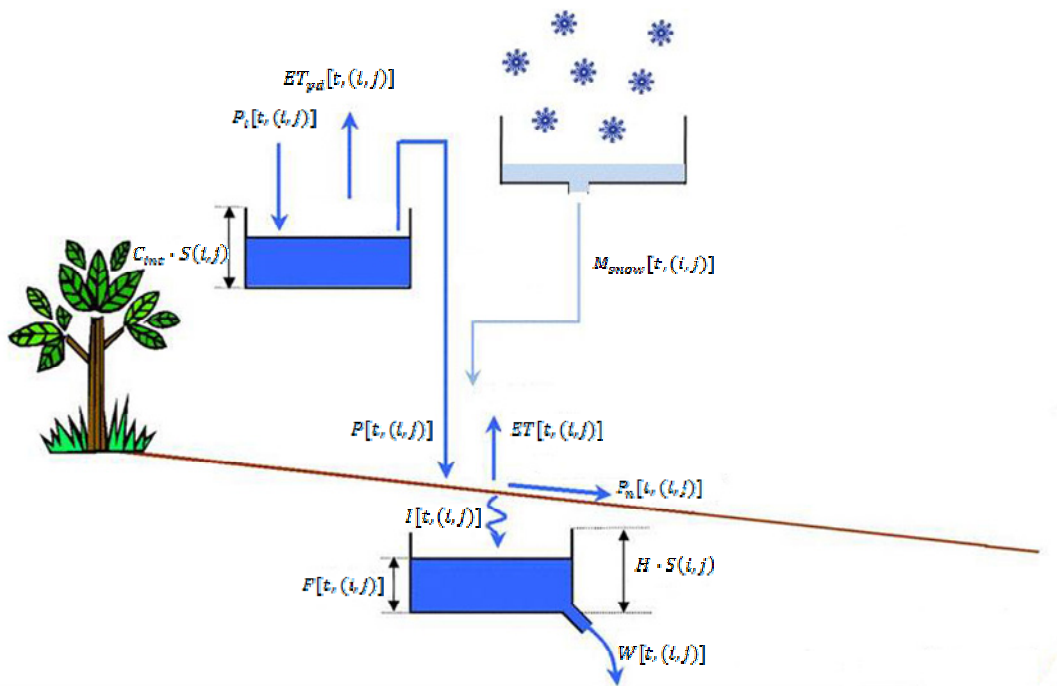
Come già detto si discretizza il bacino oggetto di studio attraverso celle elementari, di fatto coincidenti con pixels del modello digitale del terreno, definite ciascuna dalla posizione  $(i, j)$  e caratterizzate dalla medesima risoluzione spaziale. Di seguito si presenta la schematizzazione adottata per descrivere i processi che portano alla formazione di deflusso superficiale su ogni singola unità elementare.

Gli autori assumono che su ogni cella  $(i, j)$  di discretizzazione del bacino si trovi un serbatoio sotterraneo lineare (*serbatoio di infiltrazione*), che si riempie per effetto dell'infiltrazione e si svuota tramite uno scarico lineare al fondo attraverso il quale l'acqua passa dalla superficie del terreno alla falda profonda. La capacità del serbatoio di infiltrazione è una variabile che assume valore differente in corrispondenza di ogni cella ed è determinata a priori dall'utilizzatore del

modello. Si sceglie infatti di ricorrere, come nel caso del modello SCS-CN, alla rappresentazione della capacità di questo serbatoio in funzione del valore CN, che racchiude in sé considerazioni riguardanti le caratteristiche di tipo e uso di suolo oltre che di umidità preesistente. In particolare la capacità massima del serbatoio di infiltrazione localizzato sulla cella  $(i, j)$  è definita come  $H \cdot S(i, j)$ , avendo indicato con  $H$  un parametro di calibrazione adimensionale e con  $S(i, j)$  [L] la capacità di ritenzione idrica del metodo SCS-CN classico:

$$S(i, j) = 254 \left( \frac{100}{CN(i, j)} - 1 \right) \quad (1.9)$$

In Figura I- 2 si trova la rappresentazione schematica dei processi descritti dal modello AFFDEF a scala locale.



**Figura I- 2:** Rappresentazione schematica dei fenomeni di interazione tra suolo, atmosfera e vegetazione riprodotti dal modello AFFDEF in ogni cella di discretizzazione del bacino.

Se la precipitazione si presenta in forma liquida viene intercettata e invasata all'interno del serbatoio di intercettazione, la cui capacità massima è uguale  $C_{int} \cdot S(i, j)$ , essendo  $C_{int}$  un parametro di taratura adimensionale ed  $S(i, j)$  [L] la capacità di ritenzione idrica locale definita in equazione (1.9). La quantità di acqua che, sotto forma di evapotraspirazione, esce dal sistema viene anzitutto sottratta ad ogni passo temporale dal quantitativo contenuto all'interno del serbatoio di intercettazione.

Il serbatoio di intercettazione riceve quindi in ingresso la pioggia lorda  $P_l[t, (i, j)]$ , determinata attraverso il metodo dei poligoni di Thiessen o dell'inversa distanza al quadrato, che viene accumulata fino al raggiungimento, al netto dell'evapotraspirazione  $ET_{pd}[t, (i, j)]$ , della capacità massima del serbatoio stesso. La quantità di acqua in eccesso,  $P[t, (i, j)]$ , raggiunge il suolo e ad essa si somma anche la porzione dovuta allo scioglimento

$M_{snow}[t, (i, j)]$  dello strato di neve eventualmente presente sulla cella. Si indica con  $PM[t, (i, j)]$  la somma di queste due quantità:

$$PM[t, (i, j)] = P[t, (i, j)] + M_{snow}[t, (i, j)] \quad (1.10)$$

A livello del suolo è presente il secondo serbatoio, il serbatoio di infiltrazione, caratterizzato da una capacità di ritenzione massima  $H \cdot S(i, j)$ , che provvede alla separazione dei flussi superficiali da quelli profondi. La legge che regola la partizione dei deflussi è del tutto analoga a quella prevista dal metodo SCS-CN classico. La sostanziale differenza sta nel fatto che il modello del *Curve Number modificato* permette a parte del contenuto di acqua presente nel serbatoio di infiltrazione di alimentare la falda profonda. Attraverso la seguente equazione si applica la separazione dei deflussi:

$$\frac{P_n[t, (i, j)]}{PM[t, (i, j)]} = \frac{F[t, (i, j)]}{H \cdot S(i, j)} \quad (1.11)$$

nella quale  $F[t, (i, j)]$  (mm) è il contenuto di acqua presente nel serbatoio di infiltrazione al tempo  $t$  e  $P_n[t, (i, j)]$  (mm) è il deflusso superficiale che si genera al tempo  $t$  sulla cella  $(i, j)$ .

La porzione di acqua che effettivamente penetra all'interno del suolo (l'infiltrazione) è calcolata come:

$$I[t, (i, j)] = PM[t, (i, j)] - P_n[t, (i, j)] \quad (1.12)$$

nella quale tutti i contributi sono espressi in (mm).

Il deflusso dallo scarico di fondo  $W[t, (i, j)]$  che va ad alimentare la falda profonda si ottiene poi mediante la:

$$W[t, (i, j)] = \frac{F[t, (i, j)]}{H_s} \quad (1.13)$$

nella quale  $H_s$  [t] è un ulteriore parametro di calibrazione ed  $F[t, (i, j)]$  [L] l'effettivo accumulo di acqua all'interno del serbatoio di infiltrazione.

La seconda differenza sostanziale tra il modello SCS-CN classico e quello proposto da Moretti e Montanari (2007) riguarda lo schema del processo di evapotraspirazione; nel caso in cui sia presente acqua nel serbatoio di intercettazione, il contributo dell'evapotraspirazione è sottratto dal contenuto idrico di quest'ultimo. In questo caso l'evapotraspirazione effettiva è assunta pari a quella potenziale  $ET_{pd}[t, (i, j)]$  stimata con l'equazione (1.8). Qualora il serbatoio di intercettazione non contenga acqua a sufficienza, il contributo all'evapotraspirazione mancante  $ET[t, (i, j)]$  è sottratto dal contenuto del serbatoio di infiltrazione.

L'equazione che governa il bilancio idrico del serbatoio di infiltrazione si può dunque scrivere come:

$$\frac{dF[t, (i, j)]}{dt} = I[t, (i, j)] - W[t, (i, j)] \quad (1.14)$$

Esplicitando tutti i termini in gioco, evapotraspirato compreso, si ottiene l'equazione di bilancio di massa nella forma:

$$\frac{dF[t, (i, j)]}{dt} = PM[t, (i, j)] \cdot \left\{ 1 - \frac{F[t, (i, j)]}{H_s} \right\} - \frac{F[t, (i, j)]}{H_s} - ET[t, (i, j)] \quad (1.15)$$

che può essere risolta alle differenze finite.

Di fatto quindi nei periodi asciutti il modello del Curve Number modificato è in grado di simulare la diminuzione di acqua all'interno del suolo per effetto del drenaggio verso la falda profonda e dell'evapotraspirazione.

## 1.6 Propagazione dei deflussi superficiali e profondi

Con la schematizzazione descritta al paragrafo precedente si simula la generazione di deflusso superficiale e profondo in ogni cella  $(i, j)$  per ogni istante temporale  $t$ . Moretti & Montanari (2007) propongono che questi vengano propagati lungo i versanti, all'interno del reticolo idrografico e nel sottosuolo, mediante il metodo di *Muskingum-Cunge* a parametri variabili nello spazio e nel tempo (Cunge, 1969). Inoltre si assume che, nella propagazione dei deflussi superficiali, il valore dei parametri stessi sia differente per il caso di propagazione sui versanti da quello di propagazione all'interno della rete drenante.

Un approccio di questo genere si dimostra affidabile ed efficace nel caso di simulazioni in continuo a scala di bacino ed assumendo un passo temporale orario; l'obiettivo che ci si propone in questo studio è invece differente: si vuole simulare il bilancio idrologico di una vasta area, localizzata sul versante emiliano dell'Appennino. Per questo è necessario disporre di uno strumento che dia una rappresentazione accurata della variabilità spaziale dei processi che concorrono alla formazione dei deflussi, ma che al contempo sia snello e sufficientemente veloce. Per questo si è scelto di adottare un passo temporale non più orario, ma giornaliero. Per lo stesso motivo si cerca di mantenere un numero di parametri che necessitano di calibrazione sufficientemente basso. Se si adottasse il metodo di *Muskingum-Cunge* a parametri variabili anche nel caso in esame si incorrerebbe fondamentalmente in due problemi:

- la necessità di calibrare un numero di parametri effettivamente troppo elevato, rendendo AFFDEF un modello non adatto all'obiettivo che ci si è proposti,
- la violazione della condizione di Courant.

I bacini dell'Appennino emiliano-romagnolo sono infatti caratterizzati da tempi di corrivazione sensibilmente bassi; si parla infatti di bacini nei quali l'onda di piena raggiunge la sezione di chiusura in tempi paragonabili al giorno. Scegliendo un passo di integrazione, per le equazioni del modello afflussi-deflussi, superiore a 1 ora, si incorrerebbe nella mancanza di convergenza numerica della soluzione delle equazioni adottate. Supportati da queste considerazioni si è scelto di adottare uno schema semplificato per la propagazione dei deflussi.

Volendo simulare il comportamento contemporaneo di bacini caratterizzati da tempi di corrivazione paragonabili con il passo di simulazione scelto, 1 giorno, è ammissibile trasferire semplicemente i deflussi superficiali generati in ogni singola cella  $(i, j)$  direttamente alla sezione

di chiusura del bacino. La propagazione dei deflussi superficiali è quindi rappresentata dal trasferimento diretto di tutti i singoli contributi alla sezione di chiusura, in corrispondenza della quale si procede alla loro somma.

Gli scarichi al fondo di tutti i serbatoi di infiltrazione presenti sulle celle, in cui è stato discretizzato il bacino, alimentano invece un unico serbatoio lineare che vuole rappresentare il comportamento proprio della falda profonda. Questo raccoglie dunque i singoli deflussi  $W[t, (i, j)]$  e gestisce la risposta lenta del bacino rilasciando il proprio scarico al fondo in corrispondenza della sezione di chiusura tramite la:

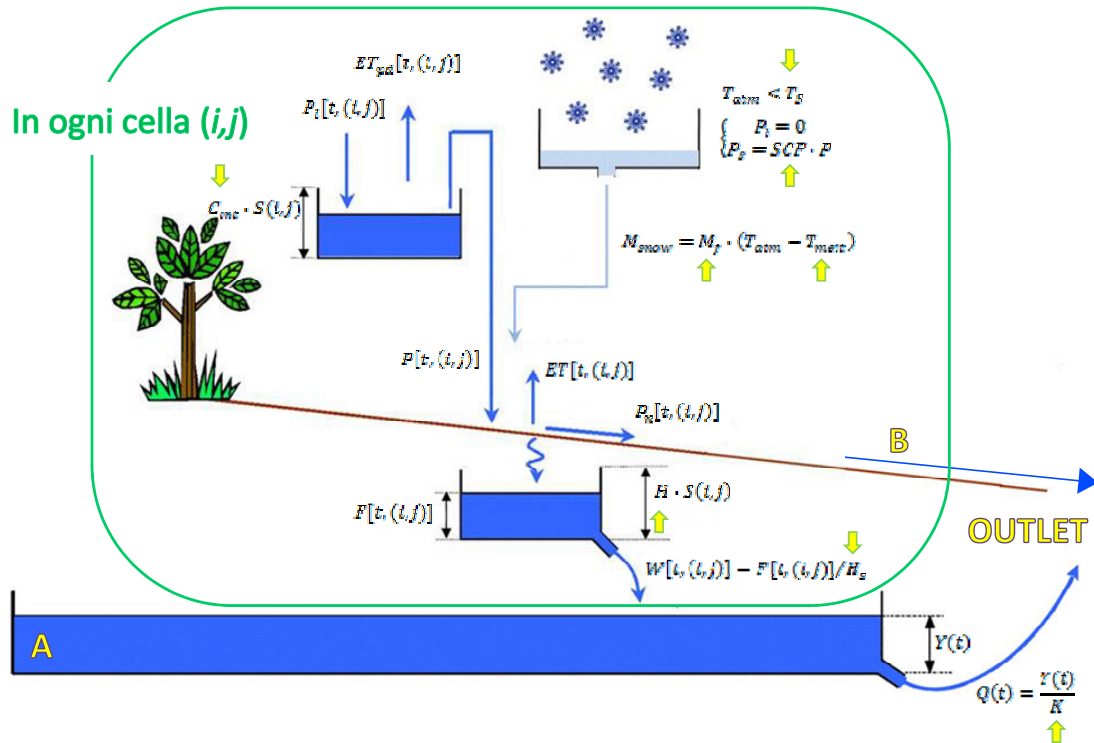
$$Q(t) = \frac{Y(t)}{K} \quad (1.16)$$

essendo  $Q(t)$  lo scarico al fondo,  $K$  [t] la costante di rilascio propria del serbatoio profondo ed  $Y(t)$  la quantità di acqua effettivamente immagazzinata all'interno del serbatoio profondo all'istante  $t$ .

Il bilancio idrico del serbatoio profondo è assicurato tramite l'equazione:

$$\frac{dY(t)}{dt} = \sum_{i,j} W[t, (i, j)] - Q(t) \quad (1.17)$$

La portata defluente in corrispondenza della sezione di chiusura del bacino è dunque data dalla somma dei due differenti contributi: da un lato la somma dei deflussi superficiali,  $P_n[t, (i, j)]$  stimati per ogni cella, moltiplicata per l'area del bacino e dall'altro il deflusso profondo  $Q(t)$ .



**Figura I- 3:** Schema di funzionamento del modello AFFDEF e dei parametri di calibrazione

In Figura I- 3 si riporta la schematizzazione utilizzata per il modello AFFDEF adattato per essere applicato al caso di studio; si distinguono in particolare all'interno del riquadro verde i processi che vengono simulati in corrispondenza di ogni unità di discretizzazione del bacino ed all'esterno dello stesso i processi di trasferimento dei deflussi che portano alla formazione di deflusso superficiale in corrispondenza della sezione di chiusura del bacino. Con la lettera B si identificano i deflussi superficiali prodotti da ogni singola cella e che vengono trasferiti direttamente nella sezione di chiusura del bacino, mentre con la lettera A si denota il serbatoio profondo di accumulo dei deflussi di falda provenienti da ogni unità fondamentale.

## 1.7 I parametri del modello AFFDEF

La descrizione delle varie componenti di AFFDEF ha più volte messo in luce la presenza di parametri, il cui valore deve essere stimato in riferimento al bacino oggetto di studio. Si distinguono in particolare:

- il parametro CN, adimensionale, stimato a priori dall'utilizzatore per ogni elemento di discretizzazione del bacino; in ingresso al modello AFFDEF viene dunque fornito, oltre al modello digitale del terreno anche una mappa contenente i valori assunti dal CN in ogni cella. In linea di principio CN può assumere valori variabili tra 0 (si riferisce ad un suolo perfettamente permeabile) e 100 (indica la presenza di un suolo completamente impermeabile), è tabellato in funzione dell'uso e del tipo di suolo oltre che del grado di potenziale assorbimento di acqua e della pioggia caduta nei 5 giorni precedenti;
- i parametri  $\alpha$  (mm) e  $\beta$  (mm°C<sup>-1</sup>) sono i coefficienti di regressione coinvolti nel processo di evapotraspirazione e stimati al di fuori della modellazione afflussi-deflussi dall'utilizzatore (ci si riferisca all'appendice A per maggiori dettagli);
- i parametri coinvolti nella generazione dei deflussi superficiali e profondi in ogni unità di discretizzazione del bacino. Questi assumono valore costante durante tutto l'intervallo di simulazione e si mantengono omogenei su tutto il bacino, cioè assumono lo stesso valore in corrispondenza di tutte le celle ( $i,j$ );
- il parametro  $K$  [t] costante di rilascio del serbatoio profondo che assume valore costante nel tempo.

I parametri appartenenti agli ultimi due gruppi, ed indicati attraverso una freccia gialla in Figura I- 3, vengono tarati in base al confronto diretto tra i deflussi osservati in corrispondenza della sezione di chiusura del bacino idrografico oggetto di studio e le portate simulate dal modello AFFDEF sulla stessa sezione.

Parametro	Descrizione	Dimensioni
$C_{int}$	Costante moltiplicativa (capacità del serbatoio di intercettazione)	[-]
$T_s$	Temperatura di soglia, al di sotto della quale le precipitazioni sono nevose	[T]
$T_{melt}$	Temperatura al di sopra della quale si ha scioglimento del manto nevoso	[T]
$M_f$	Fattore gradi-giorno	[L/T/t]
SCF	Fattore correttivo che considera la presenza di errori di misura in caso di neve	[-]
$H$	Costante moltiplicativa (capacità del serbatoio di infiltrazione)	[-]
$H_s$	Costante di rilascio del serbatoio di infiltrazione	[t]
$K$	Costante di rilascio del serbatoio profondo	[t]

**Tabella I- 1:** Elenco dei parametri del modello AFFDEF che richiedono la calibrazione



## 2 L'incertezza nei modelli

La variabilità spaziale dei parametri di modelli idrologici di tipo spazialmente distribuito è stata, fin ad ora, poco studiata, soprattutto a causa delle alte risorse computazionali che richiede il problema.

Volendo, con questo studio, gettare le basi per procedere alla stima della risorsa idrica presente in un'area di vaste dimensioni, ci si pone l'obiettivo di effettuare simulazioni di lungo periodo non solo su un unico bacino idrologico, ma su numerosi contemporaneamente. Si moltiplicano in questo modo gli sforzi richiesti per procedere alla calibrazione del modello scelto su ogni singolo bacino. La regione che si sceglie di studiare è caratterizzata da un'alternanza di zone montuose, collinari e pianeggianti, zone caratterizzate dunque da differenti regimi di precipitazione e temperatura, si necessita dunque di uno strumento che renda conto della variabilità spaziale dei fenomeni che concorrono alla formazione dei deflussi; la scelta del modello da adottare ricade dunque sui modelli spazialmente distribuiti, per la loro capacità di adattarsi con maggiore dettaglio alla variabilità dei fenomeni naturali.

La simulazione dei fenomeni che avvengono all'interno di un singolo bacino tramite un modello afflussi-deflussi spazialmente distribuito è già di per sé un problema spinoso: questo tipo di modelli solitamente coinvolge un alto numero di parametri che necessitano di una fase di calibrazione. Al contempo però questi modelli offrono un sostanziale vantaggio: spesso i parametri possono avere un significato fisico o possono essere strettamente legati a proprietà osservabili del sistema, implicando quindi, teoricamente, la possibilità di stimarne a priori il valore. Le differenze di scala, l'alto numero di parametri in gioco, la presenza di errori nelle osservazioni evita però che la stima effettuata sia accurata (*Yadav et al., 2007*). Il processo di calibrazione, dunque indispensabile in questo caso, può comunque essere semplificato identificando a priori un prefissato intervallo di validità all'interno del quale effettuare la ricerca.

Alcuni parametri possono poi avere carattere di omogeneità nello spazio, dunque presentare valori sensibilmente "simili" su aree differenti; nel momento in cui si voglia stimare il bilancio idrologico di una regione, si effettuano simulazioni idrologiche contemporaneamente su più bacini. Se è possibile dimostrare la sostanziale costanza di alcuni parametri nello spazio, è possibile ridurre il numero complessivo di parametri che necessitano della fase di calibrazione e dunque è possibile ridurre i tempi di calcolo. Nella trattazione che segue si pone l'accento proprio sulla verifica della fattibilità di questa idea cercando di verificare se sia possibile dimostrare che alcuni parametri risultano omogenei nello spazio.

La stima dei parametri del modello, così come l'adozione di modelli prestazionali sia in calibrazione che in fase di previsione sono tematiche che si complicano per la presenza di incertezza (*Montanari, 2007*); tutte le componenti proprie di un modello sono infatti affette da

incertezza. Si pensi ad esempio alle variabili in ingresso al modello, caratterizzate da errori di misura (ad esempio sono frequenti gli errori nella misura di precipitazione e temperatura), così come agli errori di misura nella serie di portate osservate, oltre all'incertezza strutturale di un modello e a quella legata ai suoi parametri. Un approccio corretto sarebbe dunque quello di considerare come un simulatore del sistema ugualmente probabile qualsiasi modello, qualsiasi combinazione di parametri che preveda la variabile o le variabili di interesse. Non c'è quindi alcuna ragione di aspettarsi che esista un unico set di parametri, e dunque un modello particolare, che riproduca al meglio la realtà fisica di un dato bacino idrologico. Piuttosto ci si dovrebbe riferire alla verosimiglianza, o alla possibilità che un insieme di valori (set di parametri) simuli in modo accettabile il comportamento del bacino stesso. Questo ragionamento si discosta dal classico approccio utilizzato in idrologia, con il quale solitamente si vuole determinare un insieme di parametri che riproduce l'ottimo globale (*Beven & Binley, 1992*). Lo studio dell'incertezza legata ai parametri, tema sul quale verteranno le analisi descritte in questa sezione, è stato affrontato il più delle volte in letteratura ricorrendo a simulazioni Monte Carlo e riferendosi quasi sempre ad un unico bacino idrografico e prendendo in esame modelli di tipo concettuale e a parametri concentrati. E' il caso ad esempio di *Gupta et al. (1998)* i quali affrontano questo tema studiando il bacino del fiume Leaf (USA) attraverso il modello SAC-SMA a 16 parametri e di *Uhlenbrook et al. (1999)* che analizzano l'incertezza dei parametri del modello HBV su un piccolo bacino montano. Un'alternativa alle simulazioni Monte Carlo per lo studio dei parametri di un modello concettuale derivato da HBV viene proposta da *Merz et al. (2009)* i quali paragonano i valori dei parametri e i valori assunti dalle efficienze del modello calibrato su differenti periodi. Affinchè quest'ultimo genere di analisi sia significativa sono necessarie lunghe serie di osservazioni di portata.

L'ipotesi di costanza dei parametri su differenti bacini risulta verificata se l'incertezza a loro associata è sufficientemente bassa e se i valori che assumono sui differenti bacini sono statisticamente compatibili.

## **2.1 La calibrazione dei modelli afflussi-deflussi**

La calibrazione di un modello afflussi-deflussi è il processo tramite il quale si selezionano i valori dei parametri in modo che il modello stesso riproduca al meglio la risposta del bacino ad una sollecitazione atmosferica. Affinchè un modello sia utile in fase di previsione, i valori dei parametri devono accuratamente rispecchiare le proprietà delle componenti del sistema che vengono schematizzate nella descrizione dei fenomeni fisici coinvolti nella produzione di portata. Di solito non è possibile procedere alla misura diretta dei parametri stessi, che quindi necessitano di essere stimati indirettamente minimizzando la discrepanza tra la risposta del bacino simulata dal modello e quella effettivamente osservata. A questo scopo si ricorre usualmente al confronto tra gli idrogrammi di portata simulati dal modello e osservati in corrispondenza della particolare sezione di chiusura del bacino. In letteratura sono numerosi gli esempi che sottolineano come la calibrazione di modelli piogge-deflussi sia particolarmente difficoltosa anche a causa dell'alta numerosità dei parametri (*Sorooshian & Gupta, 1983*). Si sottolinea infine l'importanza rivestita da una corretta calibrazione ricordando che, paradossalmente, se i parametri sono mal specificati anche il modello più elaborato, produrrà portate sempre relativamente lontane da quelle osservate, rivelandosi dunque inaffidabile.

E' possibile ricorrere ad una calibrazione dei parametri di tipo manuale oppure a calibrazioni di tipo automatico; con la prima metodologia il giudizio esperto di un idrologo, sfruttando le informazioni sul bacino, stabilisce come e di quanto variare i parametri del modello per ottenere una corretta simulazione dell'idrogramma osservato. Si tratta di una metodologia che si dimostra utile ed efficace se si calibra un modello parsimonioso (dotato di un numero limitato di parametri), poiché si riesce a distinguere quale sia l'effetto sulla portata dei diversi parametri, ma anche molto lunga e difficoltosa, nel caso di modelli dotati di numerosi parametri. Si ricorre invece alla calibrazione automatica in molte occasioni, soprattutto quando si voglia velocizzare il processo di calibrazione di un modello sofisticato o quando sia necessario disporre di un numero elevato di simulazioni.

Vrugt et al. (2008, 2009) riportano una formulazione chiara del problema generale della calibrazione, alla quale ci si riferirà spesso nel prosieguo della trattazione.

Sia  $\tilde{\mathbf{S}} = f(\theta, \mathbf{P}, \hat{\phi})$  la generica rappresentazione di un modello piogge-deflussi, nella quale si indica con  $\tilde{\mathbf{S}} = \{\tilde{s}_1, \dots, \tilde{s}_n\}$  la serie di portate simulate dal modello  $f$  il cui ingresso è costituito dai dati osservati di pioggia, temperatura ed eventualmente evapotraspirazione,  $\mathbf{P}$ , e sia caratterizzato dai parametri  $\theta$ . Sia poi  $\mathbf{S} = \{s_1, \dots, s_n\}$  il vettore degli  $n$  valori osservati di portata nella sezione di chiusura del bacino. La differenza tra i valori previsti come risposta del modello e le misure di portata può essere espressa tramite il vettore dei residui  $\mathbf{E}(\theta)$ :

$$\mathbf{E}(\theta) = \{\tilde{\mathbf{S}} - \mathbf{S}\} = \{\tilde{s}_1 - s_1, \dots, \tilde{s}_n - s_n\} = \{e_1(\theta), \dots, e_n(\theta)\} \quad (2.1)$$

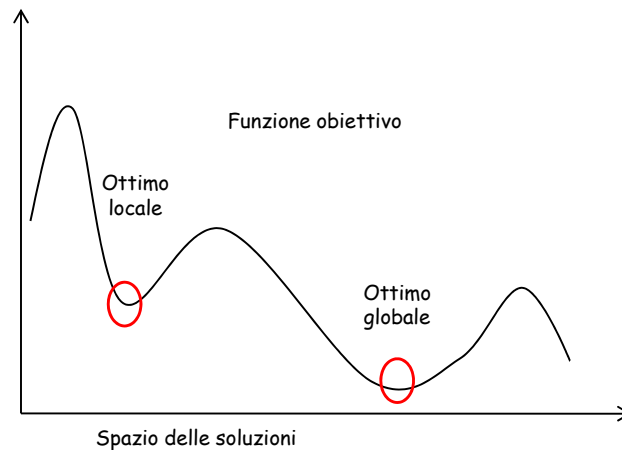
Più i residui si avvicinano allo zero e migliore è il risultato della calibrazione. La procedura di calibrazione classica punta cioè ad ottenere la minore differenza tra i valori riportati nel vettore delle portate simulate,  $\tilde{\mathbf{S}}$ , ed i valori di portata osservati,  $\mathbf{S}$ , attraverso l'ottimizzazione della funzione obiettivo scelta. Il metodo dei minimi quadrati è un approccio tradizionale e consiste nella minimizzazione della funzione  $F(\theta)$  così definita:

$$F(\theta) = \sum_{i=1}^n e_i(\theta)^2 \quad (2.2)$$

Le simulazioni così definite risentono anche della particolare scelta effettuata a riguardo delle condizioni iniziali  $\hat{\phi}$ ; si nota però che la dipendenza dei valori simulati dal modello dalla scelta di  $\hat{\phi}$  diminuisce all'aumentare della lunghezza della serie simulata stessa, cioè all'aumentare della distanza dal punto di inizio della simulazione. Per affievolire la sensibilità del modello adottato dalla scelta delle condizioni iniziali  $\hat{\phi}$  si è soliti utilizzare un *periodo di riscaldamento*, periodo durante il quale i valori simulati dal modello non vengono confrontati con quelli osservati, non andando dunque a condizionare il processo di calibrazione.

La ricerca dei valori da attribuire ai parametri del modello per ottenere una corretta calibrazione dello stesso avviene attraverso l'applicazione di procedure della ricerca dell'ottimo, o *algoritmi di ottimizzazione*. Esistono diverse classi di algoritmi, tutte accomunate dalla stessa filosofia: a partire da un'ipotesi di tentativo, cioè da un insieme di valori attribuiti dall'utilizzatore ai parametri del modello, seguendo opportune "regole", si tenta di migliorare il

valore che la funzione obiettivo assume, scegliendo di volta in volta un differente insieme dei parametri. Sono proprio le “regole” adottate dalla procedura per scegliere il set di parametri da analizzare ad ogni passo successivo a definire la differenza tra gli svariati algoritmi di ottimizzazione esistenti. Ogni algoritmo di ottimizzazione si differenzia infatti per la particolare scelta effettuata dopo il confronto tra la funzione obiettivo definita al passo (i+1) e quella definita al passo (i). Si tratta di una procedura che si ripete sempre uguale a sé stessa, iterativamente, definendo di volta in volta una nuova ipotesi di tentativo e verificandone la bontà tramite il valore assunto dalla funzione obiettivo. Un problema frequentemente riscontrato è quello dei *minimi locali*: un algoritmo che esplora lo spazio delle soluzioni solitamente comincia la ricerca nell’intorno di un particolare punto focalizzandosi sulla soluzione più promettente nell’intorno. L’obiettivo della ricerca stessa è quindi quello di identificare il punto di minimo globale della funzione obiettivo analizzata. Onde evitare però che la procedura di ricerca termini anzi tempo, interrompendo l’ottimizzazione stessa, è necessario prevedere l’introduzione di tecniche particolari che permettano di evitare la stagnazione o il blocco definitivo in un punto di minimo locale (si veda Figura I- 4).



**Figura I- 4:** Il problema dei minimi locali

L’algoritmo scelto ferma la ricerca in teoria quando si è identificato il punto nello spazio dei parametri in cui la funzione obiettivo assume il valore minimo assoluto, ma nella pratica è difficile sapere se questo obiettivo viene perseguito e si sceglie di arrestare l’ottimizzazione quando si verificano alcune condizioni particolari, enunciate dal *criterio di arresto* scelto. I tre criteri che solitamente si possono verificare si distinguono in:

- Criterio della convergenza sulla funzione obiettivo:

la procedura di ottimizzazione si arresta perché l’algoritmo non è in grado di migliorare, in uno o più tentativi il valore assunto dalla funzione obiettivo. Ciò potrebbe effettivamente significare che si è in presenza del minimo globale, ma anche che l’algoritmo è rimasto bloccato in una zona di minimo locale. In termini matematici:

$$\frac{(F_{i-k}(\theta) - F_i(\theta))}{F_i(\theta)} \leq \varepsilon_f \quad (2.3)$$

nella quale il valore  $\varepsilon_f$  è un valore sufficientemente “piccolo” definito dall’utente (solitamente  $\varepsilon_f \approx 10^{-3}$ ) che influisce sulla velocità del processo di ottimizzazione;  $F_i(\theta)$  ed  $F_{i-k}(\theta)$  sono rispettivamente il valore che la funzione obiettivo assume al passo  $i$  e al passo  $i-k$ .

- Criterio della convergenza sui parametri:

la calibrazione automatica si arresta poiché l’algoritmo non riesce a ottenere miglioramenti sensibili in una o più iterazioni, in quanto incapace di modificare sensibilmente il valore assunto dai parametri ed allo stesso tempo migliorare il valore della funzione obiettivo.

In termini matematici, alla equazione (2.3) si aggiunge la condizione:

$$\left| \frac{(\theta(j))_{i-k} - \theta(j)_i}{\theta(j)_{max} - \theta(j)_{min}} \right| \leq \varepsilon_{\theta(j)} \quad (2.4)$$

valida per ogni parametro  $\theta(j)$ .

Come nel caso precedente, ciò può indicare l’avvenuta individuazione del punto ottimale, ma anche il fatto che l’algoritmo ha raggiunto una zona in cui si ha forte interazione tra i parametri.

- Raggiungimento del numero massimo di iterazioni:

l’ultimo criterio introduce un limite massimo di tempo all’ottimizzazione, forzando la procedura all’arresto nel caso si superi il numero massimo di iterazioni consentite.

I precedenti criteri sono tutti implementati nell’algoritmo di ottimizzazione SCE-UA (*Duan et al., 1992*) che si adatterà nella trattazione svolta nella II parte di questo lavoro.

Si distingue invece dalla calibrazione, la validazione del modello stesso, una procedura mediante la quale si testa la capacità del modello, precedentemente calibrato, di riprodurre risultati. L’obiettivo dunque che ci si pone in validazione è quello di simulare un set di dati indipendenti, nel senso che l’informazione che contengono non è mai stata utilizzata durante la fase di implementazione del modello. Si tratta dunque di una procedura di simulazione di serie di portata mai analizzate fino a quel momento durante l’identificazione e la calibrazione del modello. L’attendibilità della validazione viene verificata mediante indici di bontà della stima e/o tecniche grafiche.

## 2.2 Equifinalità

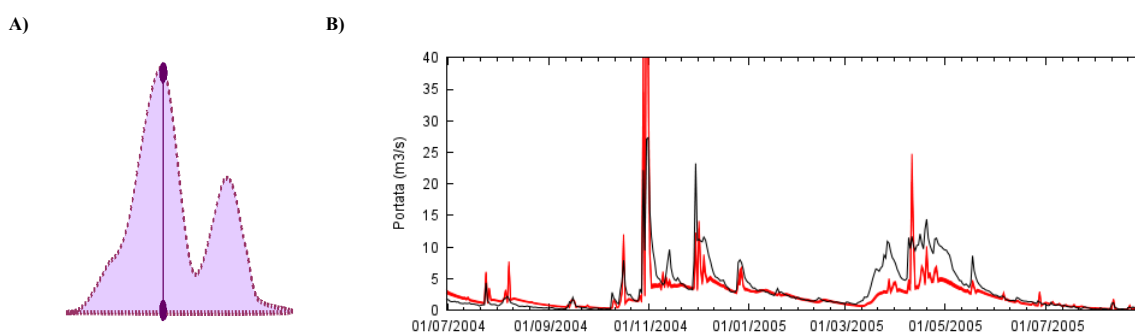
Il ricorso all’approccio classico di calibrazione appena descritto è stato spesso criticato in riferimento all’impossibilità di considerare le molteplici sorgenti di errore e fonti di incertezza coinvolte.

Nella modellazione afflussi-deflussi esistono infatti diverse fonti di incertezza: la misura delle variabili in ingresso al modello, per esempio, è affetta da incertezza, così come la definizione delle condizioni al contorno; è necessario considerare anche l’errore che si commette utilizzando un particolare modello, piuttosto che un modello di differente natura, come approssimazione dei processi che realmente si verificano (incertezza strutturale del

modello), oltre all'errore associato alle serie di portate osservate stesse (si pensi, tra le altre, all'incertezza legata alla stima delle scale di deflusso). Tutte le fonti di errore devono essere considerate esplicitamente.

La tesi dell'equifinalità (*Beven, 2006*) ha una lunga storia soprattutto in campo geomorfologico; se ci si riferisce a questo ambito è facile infatti comprendere che possono essere numerosi e di differente natura i processi che hanno contribuito alla realizzazione della particolare conformazione del suolo presente in una località; non esiste dunque un unico processo che ha portato ad ottenere quella data forma. Se si trasporta questo concetto in idrologia, allora è naturale rigettare l'ipotesi che esista, e dunque debba essere trovata, un'unica e corretta rappresentazione del sistema-bacino che porti, sollecitato da date condizioni meteo, alla formazione di un particolare regime di portate. Alcuni approcci rigettano completamente l'idea dell'esistenza di un modello "ottimale" (per un bacino specifico, caratterizzato da una particolare condizione idrologica verificatasi in un periodo ben specifico) e preferiscono invece determinare un insieme di rappresentazioni (dati in ingresso al modello, modelli dotati di strutture differenti, numerosi set di parametri) ugualmente accettabili e consistenti con le osservazioni rilevate sul sito di interesse. E' questa la base della metodologia Generalised Likelihood Uncertainty Estimation (GLUE) sviluppata da Beven and Binley (*1992*). Si tratta di un approccio con il quale è possibile considerare il concetto di equifinalità dei modelli e degli insiemi dei parametri, procedendo ad una stima dell'incertezza.

Quanto esposto sopra si traduce, in termini di idrogrammi di portata simulati, nella ricerca non più della risposta ottimale, quindi dell'idrogramma che meglio approssima la serie di portate osservate, ma nella simulazione di un consistente numero di idrogrammi, ciascuno corrispondente ad un particolare set di parametri e modelli, visualizzabili ad esempio attraverso l'utilizzo di fasce di incertezza, come presentato in Figura I- 5 B). Sul fronte dei parametri non si vuole dunque ricavare l'unico valore "ottimale" per ogni parametro stimato, ma la procedura di calibrazione restituisce un cospicuo numero di valori per ciascun parametro, analizzabili attraverso la distribuzione di densità di probabilità a posteriori del parametro considerato (si veda Figura I- 5 A)).



**Figura I- 5: A)** funzione di densità di probabilità riferita ad un generico parametro, **B)** valori delle portate osservate, in nero, e simulate, in rosso, rappresentate tramite bande di incertezza al 90%.

## 2.3 Stima dell'incertezza parametrica:

Sin qui si sono brevemente descritte le motivazioni che portano alla volontà di trattare il problema della stima dell'incertezza parametrica; di seguito si discute la metodologia usualmente adottata in letteratura.

Sia  $f$  il modello scelto per la simulazione della risposta  $\mathbf{S} = \{s_1, \dots, s_n\}$  di un bacino date le condizioni iniziali  $\hat{\phi}$  e al contorno  $\hat{\zeta}$ , e caratterizzato dal vettore dei parametri  $\theta = \{\theta_1, \dots, \theta_d\}$  di lunghezza pari a  $d$ , la cui formulazione sintetica è  $\mathbf{S} = f(\theta, \hat{\zeta}, \hat{\phi})$ . Sia poi  $\hat{\mathbf{S}} = \{\hat{s}_1, \dots, \hat{s}_n\}$  il vettore delle portate osservate alla sezione di interesse. La differenza tra  $\mathbf{S}$  e  $\hat{\mathbf{S}}$  definisce il vettore dei residui:

$$\varepsilon_i(\theta|\hat{\mathbf{S}}, \hat{\zeta}, \hat{\phi}) = s_i(\theta|\hat{\zeta}, \hat{\phi}) - \hat{s}_i \quad \text{con } i = 1, \dots, n \quad (2.5)$$

Alla luce di quanto esposto nel paragrafo precedente, l'obiettivo che ci si pone non è dunque quello di ottenere una stima dei parametri "ottimali" che descrivono il modello, ma una stima della funzione di densità di probabilità a posteriori (pdf) dei parametri  $\theta$  del modello scelto,  $p(\theta|\hat{\mathbf{S}}, \hat{\zeta}, \hat{\phi})$ .

Si può perseguire questo obiettivo facendo ricorso alla combinazione di statistica Bayesiana e tecniche di campionamento Monte Carlo; è infatti questo uno degli approcci proposti per la stima dell'incertezza (legata ai parametri, alle variabili in ingresso ai modelli, strutturale dei modelli e legata alle previsioni, nella sua versione più generale) (Vrugt et al., 2009). La metodologia GLUE, studiata da Beven and Binley (1992), la procedura URS, uniform random sampling (Uhlenbrook et al., 1999) ed il ricorso a tecniche Monte Carlo Markov Chain (MCMC) traggono dunque origine dalla questa stessa idea. Sebbene con i metodi GLUE e URS si possa esplorare dettagliatamente lo spazio dei parametri, si tratta di metodologie computazionalmente poco efficienti che necessitano di campioni molto grandi. Per questo Bates and Campbell (2001) e Vrugt et al. (2009) preferiscono piuttosto il ricorso a metodi di campionamento MCMC, seppur di diverso tipo, che generano un campione dei valori dei parametri attraverso l'accurata applicazione di catene Markov (una sola nel caso di Bates and Campbell (2001), a fronte delle numerose adottabili con l'approccio di Vrugt et al. (2009)) che convergono ad una distribuzione stazionaria.

### 2.3.1 Il teorema di Bayes:

Nel corso di questa trattazione ci si riferisce ad un generico modello piogge-deflussi  $\mathbf{S} = \{s_1, \dots, s_n\}$ , adottando la stessa simbologia e le stesse ipotesi utilizzate nel paragrafo precedente.

Si assume che gli errori  $\varepsilon_i(\theta|\hat{\mathbf{S}}, \hat{\zeta}, \hat{\phi})$  non siano correlati (siano cioè mutuamente indipendenti) e che appartengano ad una distribuzione di Gauss a varianza  $\sigma_\varepsilon^2$ . Il vettore dei parametri  $\theta$  è considerato come una variabile casuale distribuita secondo una funzione di densità di probabilità, ed è proprio tramite questa funzione che si esprime l'incertezza associata ai parametri del modello. Seguendo il teorema di Bayes la funzione di densità di probabilità a posteriori nel nostro caso si definisce come:

$$p(\theta|\hat{\mathcal{S}}) = \frac{p(\theta) \cdot p(\hat{\mathcal{S}}|\theta)}{p(\hat{\mathcal{S}})} \quad (2.6)$$

non avendo esplicitato la dipendenza da  $\hat{\zeta}, \hat{\phi}$  solo per semplicità. Con  $p(\theta)$  si denota la distribuzione a priori di  $\theta$ , la quale dunque non tiene conto di alcuna informazione riguardo ad  $\hat{\mathcal{S}}, \hat{\zeta}, \hat{\phi}$ ;  $p(\hat{\mathcal{S}})$  è invece una costante di normalizzazione, necessaria affinché la  $p(\theta|\hat{\mathcal{S}})$  sia propriamente una densità.  $p(\hat{\mathcal{S}}|\theta)$  è poi la funzione verosimiglianza del campione che assume la forma seguente:

$$p(\hat{\mathcal{S}}|\theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_e^2}} \exp\left(-\frac{(s_i(\theta|\hat{\zeta}, \hat{\phi}) - \hat{s}_i)^2}{2\sigma_e^2}\right) \quad (2.7)$$

Si massimizza la funzione di log-verosimiglianza  $l$  di equazione (2.8) per ragioni di convenienza, ricordando che gli stessi valori che massimizzano la log-verosimiglianza fanno lo stesso anche per la verosimiglianza.

$$l(\theta|\hat{\mathcal{S}}, \hat{\zeta}, \hat{\phi}) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma_e^2) - \frac{1}{2}\sigma_e^{-2} \sum_{i=1}^n (s_i(\theta|\hat{\zeta}, \hat{\phi}) - \hat{s}_i)^2 \quad (2.8)$$

Le ipotesi precedentemente citate includono anche la non correlazione degli errori, di solito non rispettata in campo idrologico, essendo spesso le serie temporali dei residui non stazionarie ed autocorrelate. Per tenere in considerazione almeno in parte l'effetto di questi errori sulla correttezza della trattazione, Vrugt et al. (2009) scelgono di ricorrere ad uno schema autoregressivo del primo ordine (AR-1) sui residui, che incorporano nella loro formulazione di funzione di log-verosimiglianza. In questo modo si riesce a considerare la presenza di errori strutturali nel modello, ma si ignorano possibili errori nelle condizioni iniziali  $\hat{\phi}$  e al contorno  $\hat{\zeta}$  che sono dunque considerate come un'approssimazione di quelle “vere”.

E' però impossibile derivare analiticamente soluzioni per l'equazione di log-verosimiglianza e dunque è necessario procedere con metodi iterativi; per questo si ricorre al campionamento Monte Carlo utilizzando catene di Markov.

### 2.3.2 L'algoritmo Random Walk Metropolis

Una catena di Markov, o processo Markoviano, è un processo stocastico nel quale la scelta dello stato successivo del sistema da analizzare, quindi la scelta del successivo set di parametri da verificare, dipende unicamente dallo stato del sistema immediatamente precedente e non da tutto il processo effettuato fino a quel punto. In particolare si segue un percorso puramente casuale con il quale, a tentativi, si spazia il dominio dei parametri spostandosi dalla posizione “vecchia”,  $\theta_{t-1}$ , della catena Markov alla “nuova”,  $\vartheta$ , seguendo le regole dettate dall'approccio scelto.

La metodologia che verrà applicata al caso di studio prevede di derivare le soluzioni della funzione di verosimiglianza tramite un particolare algoritmo MCMC, chiamato DREAM (Vrugt et al., 2008). Gli autori hanno optato per l'adozione dell'approccio di Metropolis-Hastings che si compone dei seguenti passi: a) si campiona il punto candidato  $\vartheta$  (il set di parametri da



valutare), b) si decide di accettare o rigettare il nuovo punto campionato affidandosi alla probabilità descritta dalla seguente equazione:

$$\alpha(\theta_{t-1}, \vartheta) = \begin{cases} \min\left(\frac{\pi(\vartheta)}{\pi(\theta_{t-1})}, 1\right) & \text{se } \pi(\theta_{t-1}) > 0 \\ 1 & \text{se } \pi(\theta_{t-1}) = 0 \end{cases} \quad (2.9)$$

nella quale  $\pi(*)$  denota la densità della distribuzione  $*$ .

Per comprendere meglio l'idea propria dell'approccio Metropolis-Hastings applicato a catene Markov si riporta di seguito un esempio semplice utilizzato da Bagnoli e Moroni (2006) nelle loro lezioni introduttive ai metodi Monte Carlo.

Un'applicazione del metodo Monte Carlo può essere identificata anche in questo "gioco" da praticare sulla spiaggia: si disegni sulla sabbia un quadrato ed al suo interno si iscriva un cerchio e si voglia calcolare il valore di  $\pi$  con precisione arbitraria senza usare strumenti di misura o senza ricorrere alla matematica. Più persone si collochino all'esterno del quadrato, in diverse posizioni, e lancino casualmente, e bendati, sassi verso l'interno del quadrato. Il rapporto tra il numero di sassi che cadono all'interno del cerchio e quelli che cadono all'interno del quadrato è  $\frac{\pi}{4}$ , (come il rapporto tra la lunghezza del perimetro del quadrato e della circonferenza). Poiché le persone sono rimaste immobili nelle loro posizioni durante il lancio casuale di numerosi sassi, hanno compiuto un campionamento *Monte Carlo di tipo diretto*. Se si ripete lo stesso gioco chiedendo però ad ogni persona di assumere un comportamento differente si può dare un'idea di quello che accade in un campionamento *Monte Carlo con catene di Markov*: ogni persona esplorerà lo spazio dando luogo a una differente catena Markov. Ad ognuno si chiede cioè di lanciare un sasso, sempre casualmente ed essendo bendato, verso l'interno del quadrato e poi gli si chiede di spostarsi nello stesso punto in cui è caduto il sasso appena lanciato per procedere al secondo lancio, e via di seguito. Così si spiega come la probabilità di transizione che determina il passaggio ad uno stato successivo del sistema dipenda unicamente dallo stato di sistema immediatamente precedente. Ma se un sasso cade all'esterno del quadrato, come ci si deve comportare? E' proprio a questo punto che entra in gioco il concetto di *rigetto* e l'approccio Metropolis. Se un sasso viene lanciato al di fuori del quadrato, allora la persona deve rimanere ferma sulla sua posizione, non spostarsi (poiché il set di parametri deve essere cercato sempre all'interno del dominio dei parametri) e procedere al lancio successivo sempre da quella posizione; si comprende dunque che i rigetti sono dissipatori (di ciottoli) e costosi (effettivamente si spende un tentativo, ma si rimane immobili nella stessa posizione).

La probabilità dunque che dalla posizione  $a$  ci si sposti nella posizione  $b$ ,  $P(a \rightarrow b)$ , è rappresentabile tramite l'equazione:

$$P(a \rightarrow b) = \min\left(1, \frac{\pi(b)}{\pi(a)}\right) \quad (2.10)$$

Si immagini di essere in  $a$ , cioè all'interno del quadrato, la densità di probabilità associata a questa situazione è  $\pi(a) = 1$ . Si lanci a questo punto casualmente il sasso in  $b$ ; se  $b$  si trova all'interno del quadrato, allora  $\pi(b) = 1$  e dunque si accetta lo spostamento in  $b$  con probabilità

pari a 1, se invece  $b$  è al di fuori del quadrato,  $\pi(b) = 0$ , si rigetta il movimento e la catena di Markov procede al tentativo successivo mantenendo la posizione  $a$ . Si noti come l'equazione (2.10) sia sostanzialmente una semplificazione della equazione (2.9) che descrive la condizione di rigetto implementata in DREAM.

Abbandonando la notazione adottata per descrivere il gioco dei sassi e ritornando alla descrizione del condizione di *Random Walk Metropolis* inserita nell'algoritmo DREAM, questa originariamente prevedeva anche che ad ogni passo della catena sia verificata la proprietà di equiprobabilità:

$$p(\theta_{t-1}) \cdot p(\theta_{t-1} \rightarrow \vartheta) = p(\vartheta) \cdot p(\vartheta \rightarrow \theta_{t-1}) \quad (2.11)$$

nella quale  $\theta_{t-1}$  e  $\vartheta$  indicano rispettivamente il set di parametri adottato al passo precedente e il “nuovo” candidato;  $p(\theta_{t-1})$  e  $p(\vartheta)$  indicano la probabilità che il sistema si trovi nella configurazione  $\theta_{t-1}$  o  $\vartheta$ , mentre  $p(\theta_{t-1} \rightarrow \vartheta)$  e  $p(\vartheta \rightarrow \theta_{t-1})$  le probabilità condizionate che si effettui lo spostamento dalla posizione  $\theta_{t-1}$  alla  $\vartheta$  e viceversa.

L'estensione di questo algoritmo ha portato all'implementazione della tecnica Metropolis-Hastings (*Hastings, 1970*) che prevede invece la possibilità di includere distribuzioni non simmetriche,  $q(\theta_{t-1} \rightarrow \vartheta) \neq q(\vartheta \rightarrow \theta_{t-1})$ , nelle quali cioè il “salto” dalla posizione  $\theta_{t-1}$  alla candidata  $\vartheta$  non presenti la stessa probabilità di accadimento del salto contrario.

## 2.4 L'algoritmo Differential Evolution Adaptive Metropolis (DREAM)

L'applicazione agli algoritmi MCMC della tecnica Metropolis-Hastings può portare in realtà ad una convergenza del processo fin troppo lenta. Vrugt et al (*2009a*) propongono dunque un espediente che prevede l'inserimento di aggiustamenti della direzione e della grandezza del “salto” che viene effettuato ad ogni successivo spostamento previsto all'interno delle catene di Markov sulla base di informazioni provenienti dagli stati visitati del sistema durante i passaggi precedenti. Si tratta dunque di informazioni immagazzinate durante l'evoluzione delle catene di Markov.

Nella sostanza DREAM prevede dunque la ricerca di soluzioni per l'equazione di verosimiglianza ricorrendo ad un algoritmo MCMC che fa uso dell'approccio Metropolis-Hastings opportunamente esteso in modo da riuscire a risolvere il problema anche in presenza di superfici di risposta caratterizzate da numerose soluzioni ottimali locali.

La metodologia si compone di seguenti passi:

- 1) sia  $\theta = \{\theta_1, \dots, \theta_d\}$  il vettore dei parametri del modello di lunghezza  $d$  per il quale si è stabilita una particolare distribuzione di probabilità a priori; si campiona lo spazio dei parametri individuando una popolazione iniziale  $\Theta$  di dimensione pari a  $N$  (essendo tipicamente  $N=d$  oppure  $N=2d$ ) scelta in accordo alla distribuzione di probabilità a priori.

- 2) si calcoli la densità  $\pi(\theta^i)$  per ogni punto  $\theta^i$  appartenente al campione  $\Theta$ , con  $i = 1, \dots, N$ , attraverso la trasformata antilog della funzione di log-verosimiglianza adottata  $l(\theta|\hat{\mathcal{S}}, \hat{\zeta}, \hat{\phi})$  (si veda l'equazione (2.8))

PER  $i = 1, \dots, N$  SI PROCEDE CON L'EVOLUZIONE DELLA CATENA

(si sviluppano quindi N diverse catene MCMC parallele simultaneamente)

- 3) si generi il punto candidato,  $\vartheta^i$  per la catena  $i$ ,

$$\vartheta^i = \theta^i + \gamma(\delta) \cdot \left( \sum_{j=1}^{\delta} \theta^{r(j)} - \sum_{n=1}^{\delta} \theta^{r(n)} \right) \quad (2.12)$$

nella quale:

- $\delta$  indica il numero di coppie utilizzate nella generazione del punto candidato
- $r(j)$  ed  $r(n)$  sono due valori compresi tra  $1, \dots, N$ , tali che  $r(j) \neq r(n) \neq i$
- il valore assunto da  $\gamma$  è funzione del numero di coppie  $\delta$ . In analogia con quanto previsto dal *Random Walk Metropolis*, una buona scelta per  $\gamma$  è  $\gamma = 2.38 / \sqrt{2 \cdot \delta \cdot d_{eff}}$ , con  $d_{eff} = d$  in prima istanza, ma che di passo in

passo può assumere valori inferiori (come conseguenza, ad esempio, di quanto accade in equazione (2.13)).

Il punto candidato  $\vartheta^i$  della catena  $i$  si scosta dal set di parametri determinati al passo precedente  $\theta^i$ , sempre nella stessa catena  $i$ , perché ottenuto attraverso uno spostamento la cui direzione ed il cui passo sono stati influenzati anche dalle informazioni provenienti da altre catene (è questo dunque il significato dell'introduzione di  $\theta^{r(j)}$  e  $\theta^{r(n)}$ ).

- 4) si sostituisca ogni elemento  $\vartheta_k^i$  del set di parametri proposto  $\vartheta^i = \{\vartheta_1^i, \dots, \vartheta_k^i, \dots, \vartheta_d^i\}$  con  $\theta_k^i$ , ricorrendo ad uno schema binomiale nel quale CR è la probabilità di crossover

$$\vartheta_k^i = \begin{cases} \vartheta_k^i, & \text{se } U \leq 1 - CR, \\ \theta_k^i, & \text{altrimenti} \end{cases} \quad d_{eff} = d_{eff} - 1 \quad (2.13)$$

con  $k = 1 \dots d$  ed  $U \in [0,1]$ , campionato da una distribuzione uniforme.

Si ricorre al crossover, alla combinazione cioè di elementi propri del set di parametri proposto  $\vartheta^i$  e di elementi propri del set di parametri analizzato al passo precedente  $\theta^i$ , per produrre un nuovo “individuo figlio” che eredita caratteristiche da entrambi i set. In questo modo è perciò possibile indurre ogni catena costantemente verso nuove direzioni anche al di fuori del sottospazio analizzato in quel momento (*Vrugt et al, 2009a*).

- 5) si calcoli la densità  $\pi(\vartheta^i)$  (come già fatto al passo 2)) e si decida se rifiutare o accettare questo candidato in accordo a quanto stabilito dalla probabilità di Metropolis:

$$\alpha(\vartheta^i, \theta^i) = \begin{cases} \min\left(\frac{\pi(\vartheta^i)}{\pi(\theta^i)}, 1\right), & \pi(\theta^i) > 0 \\ 1, & \pi(\theta^i) = 0 \end{cases} \quad (2.14)$$

- 6) se il punto candidato non viene rifiutato, allora la catena si sposta in quel punto e  $\theta^i = \vartheta^i$ , altrimenti la catena rimane ferma nella “vecchia” posizione

FINE. SI ARRESTA L’EVOLUZIONE DELLA CATENA

- 7) si rimuovono gli outliers che deteriorerebbero significativamente la bontà della procedura vista
- 8) si calcola il coefficiente di Gelman and Rubin (1992),  $\hat{R}$  che diagnostica l’avvenuta convergenza o meno di ciascuna catena, utilizzando per lo meno un campione contenente l’ultimo 50% dei set di parametri analizzati in ognuna delle  $d$  catene
- 9) se  $\hat{R} \leq 1.2$  in tutte le catene, allora si può dire che il processo è giunto a convergenza e dunque ci si ferma, altrimenti si continua con l’evoluzione delle catene.

Il metodo adottato sceglie dunque una popolazione di punti iniziale dalla quale cominciare a campionare lo spazio delle potenziali soluzioni; ciascuno di questi punti si comporta come l’inizio di una differente catena Markov. L’adozione di  $N$  catene individuali permette di campionare con maggiore facilità lo spazio delle soluzioni anche in presenza di molteplici regioni ad alto potenziale attrattivo.

DREAM prevede anche uno scambio continuo di informazioni tra le differenti catene per consentire a ciascuna di queste di apprendere nozioni utili dalle altre; ad ogni passo infatti gli elementi di  $\Theta$  contengono le informazioni maggiormente rilevanti riguardo il processo di ricerca effettuato su ogni singola catena. Inoltre, come discusso al passo 3) della procedura, la direzione e la lunghezza del passo effettuato per il progresso di ogni catena risente di informazioni derivanti dalle altre. Questo scambio di nozioni aumenta la possibilità di “sopravvivenza” di ciascuna catena, diminuendo così il numero di outliers. E’ previsto che tutta la procedura necessiti di un periodo iniziale di “riscaldamento”, durante il quale i tentativi di campionamento effettuati non appartengono ancora alle aree di maggiore attrazione del dominio dei parametri; è necessario dunque provvedere all’analisi dei risultati utilizzando i campionamenti che non competono a questo periodo. La convergenza di ciascuna catena viene poi monitorata attraverso la statistica  $R$  di Gelman and Rubin.

La funzione di densità di probabilità a posteriori derivata utilizzando DREAM contiene le informazioni richieste per sintetizzare l’effetto dell’incertezza parametrica sulla risposta del modello. Si considerino infatti tutti i set di parametri, determinati una volta terminato il periodo di “riscaldamento” della procedura e si valuti la risposta  $\mathbf{S}_k = f(\theta_k, \hat{\zeta}, \hat{\phi})$  del modello sollecitato da ciascun insieme di parametri  $\theta_k$  (con  $k = 1, \dots, J$  essendo  $J$  il numero di

realizzazioni). I risultati contenuti in questi  $J$  diversi idrogrammi di portata simulata possono essere ad esempio sintetizzati ricorrendo alla rappresentazione di fasce di incertezza al 90%.

## **3 Caso di studio:**

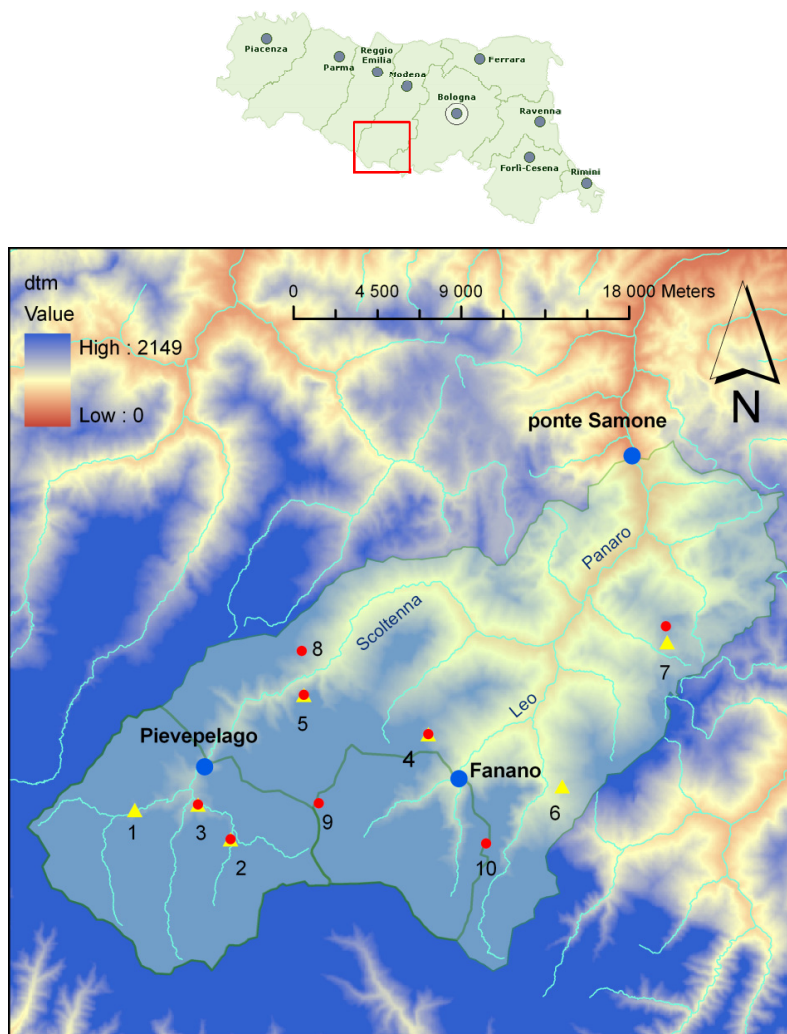
### **Stima dei parametri del modello AFFDEF**

#### **3.1 Area di studio**

Per lo studio della stima dell'incertezza parametrica si è scelto di far riferimento ad un'area che ricade all'interno del bacino del fiume Panaro, indicata in Figura I- 6. Con un'estensione pari a 584 km<sup>2</sup> il bacino del fiume Panaro occupa un'area caratterizzata da una forte escursione altimetrica; si passa infatti dai 200 m s.l.m. della sezione di chiusura posta più a valle e denominata ponte Samone, a quote superiori ai 2000 m s.l.m. dei monti che appartengono allo spartiacque appenninico, primo tra tutti il monte Cimone, sui quali solitamente sono abbondanti le precipitazioni a carattere nevoso durante l'inverno. Il fiume Panaro nasce dalla confluenza, a quote ormai proprie della media collina, dei torrenti Leo e Scoltenna, che hanno una direzione di scorrimento da Sud-Ovest verso Nord-Est, e si immette nel fiume Po 107 km più a valle rispetto alla sezione di ponte Samone.

Le precipitazioni medie della zona di alto Appennino sono abbondanti e possono raggiungere i 2200 mm annui, dei quali mediamente 1200 mm registrati nelle stagioni autunnale e invernale, mentre alla confluenza dei torrenti Leo e Scoltenna le precipitazioni medie annue si attestano su valori prossimi ai 900 mm. Un ragionamento opposto vale per le temperature medie annue, con valori che si aggirano attorno ai 2°C in alta montagna e 11 °C nella zona collinare. Il regime dei deflussi risulta torrentizio essendo caratterizzato da magre anche importanti nel periodo tardo-estivo inizio-autunnale, copiose e abbondanti morbide primaverili e piene autunnali. Poiché la porzione iniziale del bacino si trova ad alta quota, il regime delle portate è largamente influenzato dallo scioglimento primaverile delle nevi che si sono accumulate durante l'inverno.

Si sono considerate le sole stazioni idrometriche di Pievepelago (sul torrente Scoltenna), Fanano (sul torrente Leo) e ponte Samone (sul fiume Panaro) per le quali sono disponibili misure di portata giornaliera per un arco temporale di almeno 5 anni. Dopo un'analisi preliminare ed alcuni tentativi di calibrazione si è scelto invece di non prendere in considerazione il bacino del torrente Acquicciola sotteso alla sezione di Fiumalbo poiché, sviluppandosi proprio a ridosso del crinale Appenninico ed essendo di dimensioni molto ridotte, 18 km<sup>2</sup>, ha un regime estremamente difficile da simulare essendo in larga parte influenzato dai processi di accumulo e scioglimento nivale. E' dunque possibile identificare tre differenti bacini idrografici, la cui posizione e contorno è riportata in Figura I- 6, sui quali si procede alla parametrizzazione del modello afflussi-deflussi ed allo studio della stima di incertezza parametrica.



**Figura I- 6:** Modello digitale del terreno (80 m.) dell'area di studio: il bacino del fiume Panaro. In mappa sono indicate anche le posizioni delle stazioni idrometriche (cerchi blu), termometriche (cerchi rossi) e pluviometriche (triangoli gialli).

### 3.2 Scelta del modello:

Si sceglie di adottare il modello spazialmente distribuito AFFDEF (descritto al capitolo 1) e, per analizzare l'incertezza legata a ciascun parametro, si procede alla sua calibrazione su ogni bacino separatamente affidandosi all'algoritmo DREAM. Avendo la necessità di ripetere un numero elevato di simulazioni, si cerca di ridurre per quanto possibile i tempi di calcolo adottando un passo temporale pari a 1 giorno e discretizzando il bacino oggetto di studio in unità fondamentali di risoluzione spaziale pari a 1000 m, convinti così di mantenere un buon compromesso tra la volontà di descrivere la variabilità spaziale dei fenomeni responsabili della formazione del deflusso e la necessità di contenere i tempi computazionali.

Per la corretta simulazione dei deflussi il modello AFFDEF necessita di alcuni dati in ingresso:

- Matrice del modello digitale del terreno dell'area di interesse

- Mappa dei valori del Curve Number
- Osservazioni di portata in corrispondenza delle sezioni di chiusura dei bacini
- Osservazioni di pioggia e temperatura in diverse stazioni meteorologiche

Si procede dunque alla descrizione dei dati necessari per effettuare le simulazioni.

### 3.3 Informazione disponibile

Per ottenere una mappa che descriva la variabilità nello spazio del valore del curve number (d'ora innanzi definito come CN), è necessario ricorrere alla combinazione di informazioni che descrivano l'uso del suolo e il tipo di suolo. Per questo è utile riferirsi ad una mappa derivante dal progetto Corine Land Cover 2000 che contiene la descrizione della copertura del suolo riferita all'anno 2000 ed una mappa che descrive le proprietà geologiche (in particolare la litologia) dei suoli della regione Emilia Romagna (fonte R.E.R.). Le informazioni sulla litologia permettono la distinzione di 4 classi A, B, C e D, in ordine crescente di potenziale di scorrimento superficiale di ciascuna formazione e dunque anche in ordine crescente di valore di CN a parità di uso del suolo.

Pur disponendo di un modello digitale del terreno la cui risoluzione originale è pari a 80 m (fonte DICAM) e di informazioni vettoriali di dettaglio per determinare il CN, si è scelto di trasformare tutte queste informazioni in modo da poter adottare unità di discretizzazione del bacino quadrate di risoluzione pari a 1000 m. Valori alti di CN rispecchiano una bassa attitudine all'infiltrazione e, quindi, un alto grado di scorrimento superficiale.

#### 3.3.1 Informazioni idro-climatiche

Gli annali idrologici (fonte R.E.R.) riportano le misure idrometriche giornaliere per le stazioni di Pievepelago e Fanano dal 01/01/2003 al 31/12/2010, mentre nella stazione di ponte Samone le misure di portata cominciano solo il 01/01/2004. Alla data in cui si sono condotte le simulazioni erano però disponibili dati solo fino al 31/12/2008 ed è per questo che i risultati qui riportati si riferiscono unicamente al periodo compreso tra il 2003 e il 2008.

A completare poi il quadro dei dati disponibili c'è la disponibilità di misure giornaliere di temperatura media (°C) e di altezza di precipitazione cumulata giornaliera (mm) nelle stazioni meteo-climatiche indicate in Figura I- 6 (Fonte R.E.R.). Si sono rappresentate, e considerate per effettuare le simulazioni, solo le stazioni che, pur non con piena continuità, dispongono di misure giornaliere che coprono l'intervallo temporale sul quale si sono concentrate le simulazioni idrologiche (01/01/2003-31/12/2008). Si riportano in Tabella I- 2 le denominazioni delle stazioni di misura menzionate, il tipo di dato riportato (T = temperatura media giornaliera, P = altezza di precipitazione cumulata giornaliera) e, in riferimento alla temperatura, l'anno in cui la stazione si è resa operativa. Quest'ultimo dato risulta fondamentale per la stima dei valori di evapotraspirazione, calcolata con il *Radiation Method* (Todini, 1996), che prevede la stima di due parametri,  $\alpha$  e  $\beta$ , da effettuarsi riferendosi unicamente a misure storiche di valori medi annui di temperatura media mensile (per maggiori dettagli si rimanda all'appendice A).



	Denominazione	Pioggia	Temperatura	Inizio osservazioni
1	Pievepelago	P		
2	Doccia Fiumalbo	P	T	2003
3	S. Michele centrale	P	T	1959
4	Sestola	P	T	1871
5	Strettara centrale	P	T	1966
6	Rocca Corneta	P		
7	Montese	P	T	1960
8	Sassostorno		T	2003
9	Monte Cimone		T	1961
10	Lago Pratignano		T	2003

**Tabella I- 2:** Stazioni pluviometriche e termometriche utili allo studio. La numerazione riportata si riferisce alla collocazione delle stazioni riportate in Figura I- 6.

### 3.4 Stima dell'incertezza

L'incertezza parametrica del modello AFFDEF viene stimata indipendentemente per ciascun bacino, ripetendo in tutti e tre i casi la procedura nel medesimo modo. Si sceglie di adottare la metodologia messa a punto da Vrugt et al. (2008) denominata DREAM che, attraverso la combinazione di statistica Bayesiana e campionamento MCMC, porta alla stima della incertezza associata ai parametri, rappresentata attraverso la funzione di densità di probabilità a posteriori di ciascun parametro. Considerazioni sulla forma di ciascuna funzione di densità consentono di comprendere il grado di incertezza legato al parametro in considerazione; ad esempio, quanto più la particolare distribuzione di densità risulta ampiamente distribuita all'interno del dominio di variabilità, tanto maggiore sarà l'incertezza associata a quel parametro. E' possibile inoltre dedurre importanti considerazioni dal paragone tra funzioni di densità riferite al medesimo parametro ma ottenute dalla calibrazione del modello su bacini differenti.

Ci si riferisce, d'ora innanzi, ai bacini dei torrenti Leo chiuso alla sezione di Fanano e Scoltenna chiuso alla sezione di Pievepelago, definendoli rispettivamente i bacini del torrente Leo e del torrente Scoltenna, così come per il fiume Panaro si assume che l'unica sezione di chiusura, di interesse per questo studio, sia quella di ponte Samone.

Si decide di distinguere un periodo di calibrazione di lunghezza pari a 4 anni e di considerare le osservazioni di portata rimanenti per scopi di verifica della bontà di previsione del modello; poiché i torrenti Leo e Scoltenna dispongono di misure di portata giornaliere per un periodo che si estende dal 01/01/2003 al 31/12/2008, è possibile identificare gli anni 2007 e 2008 come il periodo di validazione del modello. Il fiume Panaro non è stato invece monitorato durante il 2003 e per questo motivo si è deciso di utilizzare gli anni 2004-2007 per la calibrazione del modello e l'anno 2008 viene considerato allo scopo di previsione.

Nelle sue prime battute la simulazione idrologica è altamente influenzata dalla scelta delle condizioni iniziali che l'utilizzatore del modello deve effettuare oculatamente; contenuto iniziale di umidità del suolo, di acqua all'interno del serbatoio di intercettazione o di precipitazione all'interno del serbatoio neve, ad esempio, sono quantità che effettivamente necessitano una particolare attenzione e la cui errata scelta può produrre gravi scostamenti della simulazione dal comportamento effettivo del bacino. Per svincolarsi da questo genere di problemi si è soliti considerare un periodo iniziale, di lunghezza variabile, come periodo di riscaldamento del modello, durante il quale cioè si permette al modello di adeguarsi alle

particolari condizioni del bacino. Nella pratica si considera il periodo di riscaldamento come un intervallo di tempo durante il quale il risultato della simulazione non viene utilizzato per calcolare il valore della funzione obiettivo, non andando quindi ad influenzare l'operazione di calibrazione che si sta effettuando. Nel caso che si va ora a presentare si destinano i primi 300 giorni del periodo di calibrazione al riscaldamento del modello, trascurando così, nella stima della bontà della calibrazione, il primo periodo di magra. La calibrazione vera e propria comincia dunque alla fine del mese di Ottobre del primo anno così da considerare per intero i fenomeni che portano alla formazione delle morbide e delle piene autunnali, e da includere completamente il periodo caratterizzato da precipitazioni a carattere nevoso.

La calibrazione del modello sui tre bacini è stata realizzata applicando l'algoritmo DREAM, avendo avuto cura di operare alcune scelte:

- Essendo  $d = 8$  la dimensione del problema, essendo cioè 8 i parametri che necessitano la calibrazione in AFFDEF, ci si è attenuti alle indicazioni dell'autore del modello adottando un numero di poco superiore a questo per definire la quantità di catene MCMC con le quali campionare lo spazio dei parametri. Si adottando quindi  $N = 10$  differenti catene;
- Il numero massimo di tentativi di simulazione effettuati viene definito pari a 80000, un numero cioè sufficientemente elevato da ipotizzare che l'algoritmo si fermi solo una volta arrivato a convergenza;
- Si sceglie il test definito IQR e già preimpostato come test per la verifica di possibili outlier;
- Tutti i parametri hanno una distribuzione di probabilità a priori uniforme;
- Si stabilisce che le soluzioni del problema debbano essere cercate unicamente all'interno del dominio dei parametri;
- Avendo particolare interesse per la corretta riproduzione del comportamento del bacino durante i periodi di magra, si fa in modo che l'algoritmo analizzi le portate log-trasformate, in modo così da evitare che i picchi di portata abbiano maggior peso nella simulazione; si utilizza una delle funzioni di log-verosimiglianza già presenti tra le possibili scelte offerte dall'algoritmo, in particolare si adotta una funzione definita come  $-0.5 \cdot SSR$ , essendo  $SSR$  una funzione strettamente legata alla somma degli errori al quadrato;
- Si adottano poi gli altri elementi caratteristici di DREAM in accordo a quanto già utilizzato in Vrugt et al. (2008).

Uno stratagemma suggerito poi dell'autore di DREAM per evitare che l'algoritmo incappi in errore e per rendere dunque più snello il processo di calibrazione sta nell'evitare che i parametri del modello oggetto di calibrazione siano caratterizzati da ordini di grandezza estremamente differenti. Accade infatti che degli otto parametri del modello AFFDEF, tre sono adimensionali e rappresentano delle capacità ( $C_{int}$ ,  $H$ ,  $SCF$ ) e quindi variano tra 0 ed 1, due rappresentano delle temperature di soglia ( $T_s$ ,  $T_{melt}$ ), dunque il loro ordine di grandezza si attesta tra  $10^0$  e  $10^1$ , gli ulteriori tre parametri sono invece legati all'unità di tempo e di lunghezza; l'unità di tempo adottata da AFFDEF è il secondo e quella di lunghezza è invece il mm. Se si mantengono

dunque inalterati questi parametri si ha, ad esempio, una costante  $K$  di rilascio del serbatoio profondo che, espressa in secondi, raggiunge un ordine di grandezza pari a  $10^6$ , la costante  $H_s$  di rilascio del serbatoio di infiltrazione di un solo ordine di grandezza inferiore rispetto a quella precedente ed il coefficiente gradi-giorno  $M_f$  che, espresso in  $\text{mm}/^\circ\text{C}/\text{s}$  ha un ordine di grandezza pari a  $10^{-5}$ . Il consiglio è dunque quello di log-trasformare i parametri, per lo meno quelli di ordine molto alto e molto basso ( $K, H_s, M_f$  nel caso in questione), per permettere la ricerca delle soluzioni in uno spazio dei parametri di più semplice trattazione. Si riportano per questo in Tabella I- 3 gli intervalli di variabilità dei parametri definiti anche in maniera log-trasformata, descritti in ( ).

Parametro	Descrizione	Minimo	Massimo	
$C_{int}$	Capacità del serbatoio di intercettazione	0.0005	0.80	(-)
$T_s$	Temperatura di soglia (precipitazioni nevose)	-4.5	0.1	( $^\circ\text{C}$ )
$T_{melt}$	Temperatura di soglia (scioglimento del manto nevoso)	-0.8	3	( $^\circ\text{C}$ )
$M_f$	Fattore gradi-giorno	0.0273 (-6.5)	17.239 (-3.7)	( $\text{mm}/^\circ\text{C}/\text{d}$ )
SCF	Fattore correttivo	0.9	3	(-)
$H$	Capacità del serbatoio di infiltrazione	0.0005	0.99	(-)
$H_s$	Costante di rilascio del serbatoio di infiltrazione	0.159 (2.76)	149.1 (5.73)	(h)
$K$	Costante di rilascio del serbatoio profondo	149.1 (5.73)	8784 (7.5)	(h)

**Tabella I- 3:** intervalli di validità dei parametri. ( $M_f$  in  $\log_{10}(\text{mm}/^\circ\text{C}/\text{s})$ ;  $H_s$  e  $K$  in  $\log_{10}(\text{s})$  come utilizzati dal modello)

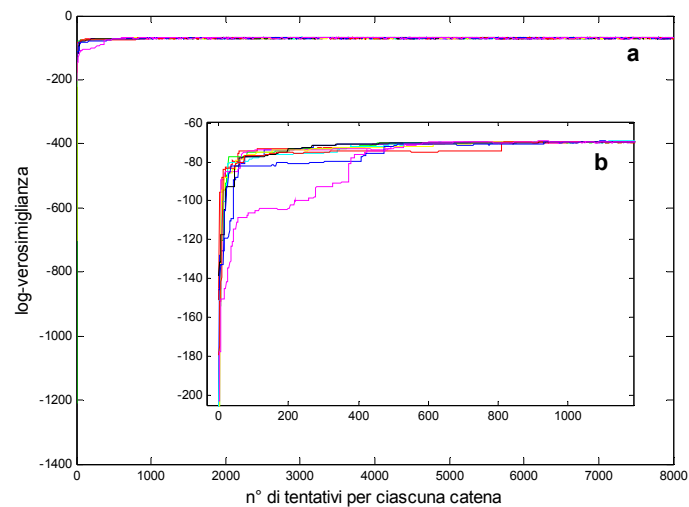
Il primo 60% dei campioni analizzati da ciascuna catena MCMC viene poi trascurato e considerato come periodo di riscaldamento dell'algoritmo DREAM (come effettuato da Vrugt et al. (2008)); in questo modo ci si mette a favore di sicurezza e le analisi di incertezza dei parametri sono condotte unicamente sui campioni delle catene Markov che già sono prossimi alla convergenza.

Si presentano di seguito i passi effettuati per assicurarsi che le simulazioni effettuate con DREAM siano effettivamente andate a buon fine e si segue, a titolo di esempio, il caso del torrente Leo.

### 3.4.1 Analisi dei risultati delle simulazioni: il caso del torrente Leo

Una volta terminata la simulazione è bene assicurarsi che il processo sia stato in grado di campionare in modo adeguato lo spazio dei parametri, indirizzando ciascuna catena Markoviana verso un'unica regione dello spazio. Per verificare ciò si può controllare che il valore assunto dalla statistica  $R^2$  di Gelman-Rubin, calcolata sull'ultimo 50% dei campioni di ogni singola catena, sia inferiore a 1.2; nel caso considerato la statistica  $R^2$  assume valori che variano tra un minimo di 1.0046 e un massimo di 1.0334, quindi il test risulta soddisfatto, indicando così che i tentativi compiuti dalle differenti MCMC portano a soluzioni che si trovano nella stessa regione di attrazione.

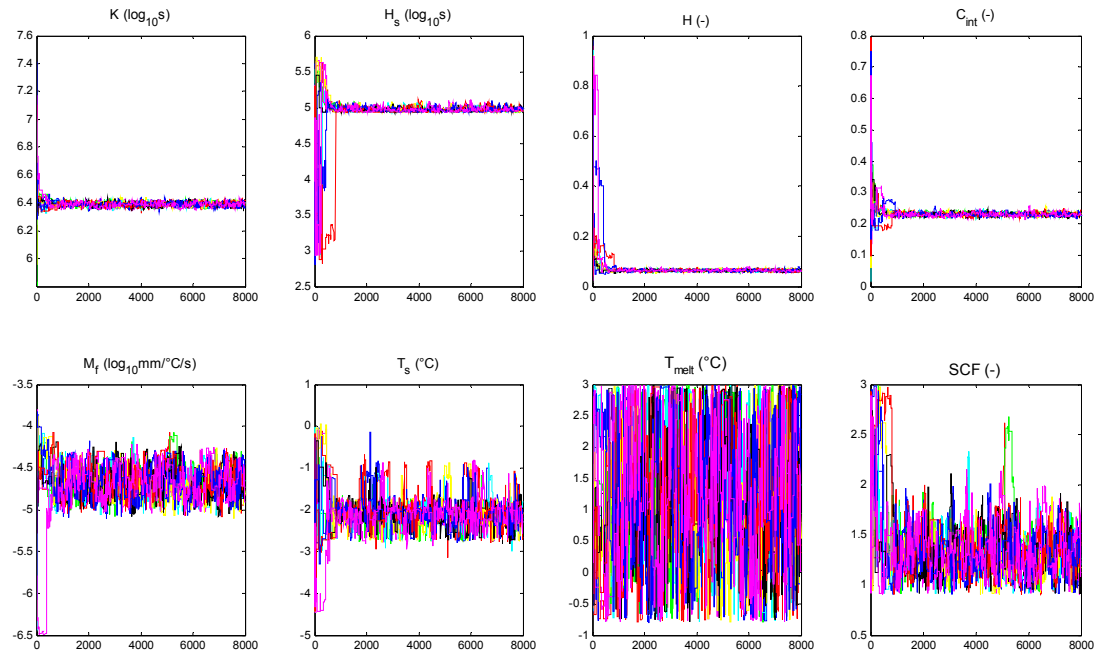
E' possibile però controllare anche con metodi grafici la convergenza di tutte le catene; in Figura I- 7 sono riportati i valori che la funzione log-verosimiglianza assume ad ogni passo, in ciascuna delle 10 diverse catene. Si noti come, dopo soli 900 tentativi compiuti da ciascuna catena Markoviana la log-verosimiglianza assuma sostanzialmente valori compresi in un intervallo estremamente ridotto (ci si riferisca in particolare al riquadro **b**). Questo indica che, ad essere rigorosi, il processo avrebbe già raggiunto sostanzialmente la convergenza dopo i primi 900 campionamenti per catena (quindi dopo un totale di 9000 tentativi) e si potrebbe dunque decidere di considerare un periodo di riscaldamento dell'algoritmo molto più piccolo rispetto al primo 60% di ogni processo. Si sceglie ad ogni modo, cautelativamente, di mantenere inalterato il valore precedentemente scelto.



**Figura I- 7:** Evoluzione della funzione log-verosimiglianza per ogni differente catena Markov (ad ogni colore corrisponde una MCMC diversa per un totale di 10 catene)

Per attestare che il processo è arrivato a convergenza non è però sufficiente analizzare l'evoluzione della funzione di log-verosimiglianza sulle diverse catene MCMC; è necessario infatti completare il ragionamento valutando anche il comportamento dei singoli parametri. Dopo un primo periodo di oscillazione casuale all'interno dell'intero intervallo di validità dei valori, affinché si verifichi la convergenza del processo è necessario che le soluzioni tentate su tutte le catene per il parametro in questione appartengano ad un stesso intervallo di valori ben definito. Riferendosi all'esempio del bacino del torrente Leo riportato in Figura I- 8, si vede come la ricerca dei valori che il parametro  $H$  può assumere spazia effettivamente su tutto l'intervallo di validità stabilito a priori (si veda quanto riportato in Tabella I- 3 ) nella prima fase di ricerca, mentre dopo i primi 1000 tentativi, la ricerca di soluzioni viene concentrata da tutte le catene all'interno di un ristretto campo. Il parametro  $T_{melt}$  invece presenta un comportamento differente, assumendo valori sempre all'interno di tutto lo spazio di validità definito a priori. L'importante, ai fini del raggiungimento della convergenza del processo è che tutte le catene Monte Carlo, riferite all'analisi del parametro in questione, presentino il medesimo comportamento. Non si può dunque che accettare quanto accade al parametro  $T_{melt}$  sottolineando però semplicemente che questo è affetto da un alto grado di incertezza. Differente invece è il caso del parametro  $SCF$ ; si nota infatti che la catena rappresentata in verde in Figura I- 7 cerca soluzioni al di fuori dello spazio delle soluzioni analizzato da tutte le altre catene. Si

tratta solo però di un caso isolato, che non perdura a lungo e, a partire da circa 5500 tentativi effettuati su ciascuna catena, lo spazio all'interno del quale vengono analizzate le soluzioni si mantiene lo stesso per tutte le catene. Per questo dunque si può continuare l'elaborazione dei risultati facendo attenzione a considerare, per la stima dell'incertezza parametrica, unicamente l'ultima frazione dei campionamenti effettuati. Si conferma così la scelta cautelativa di scartare il primo 60% delle simulazioni effettuate.

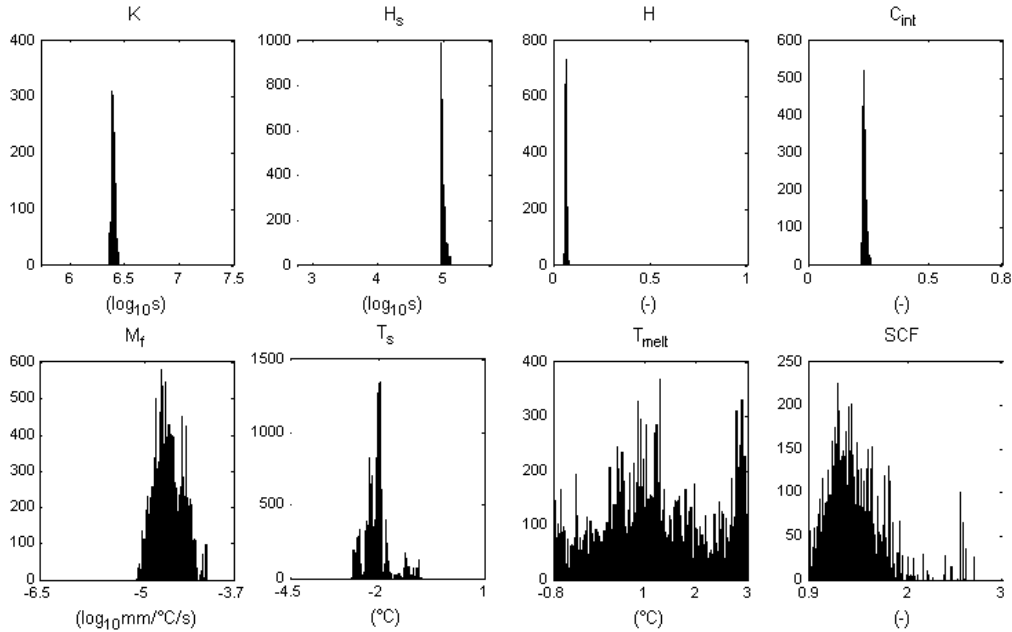


**Figura I- 8:** Valori assunti dai parametri del modello AFFDEF all'interno di ogni differente catena Markov (ad ogni colore corrisponde una MCMC diversa per un totale di 10 catene). Le unità di misura sono quelle effettivamente adottate durante la calibrazione del modello.

Giustificata così la scelta di trascurare la prima parte delle simulazioni al fine di ottenere una stima corretta dell'incertezza parametrica, si analizzino ora i risultati veri e propri.

Se si analizza la numerosità campionaria delle soluzioni determinate dal processo di calibrazione per ciascun parametro si ottiene quanto presentato in Figura I- 9; si è scelto in questo caso di mantenere le stesse unità di misura adottate in fase di calibrazione, così da poter agevolmente paragonare i risultati ottenuti con i campi di validità definiti per i parametri in Tabella I- 3 e riportati anche in figura. Si noti come nella maggior parte dei casi le distribuzioni dei parametri risultino ben definite in una regione limitata dello spazio dei parametri; si tratta di un risultato che conferma quanto già visto in Figura I- 8. In particolare questo accade per i 4 parametri che controllano la quantità di pioggia intercettata dalla vegetazione, che governano la separazione tra deflussi superficiali e profondi e che condizionano la velocità dei processi di falda. Se si analizzano invece i risultati ottenuti per i parametri che sono legati ai processi di accumulo e scioglimento nivale si nota come le incertezze in gioco aumentino; la maggiore dispersione delle soluzioni, trovate con l'algoritmo descritto, sul dominio dei parametri può portare sostanzialmente a due differenti conclusioni: da un lato può significare che si è adottato un campo di variabilità dei parametri piuttosto ristretto, mentre d'altro canto può entrare in gioco il fattore incertezza. Infatti tanto più le soluzioni presentano distribuzioni ampie, quanto

più i parametri risultano incerti. E' quanto accade al parametro  $T_{melt}$ , che governa il fenomeno dello scioglimento nivale, e che assume valori su tutto il dominio di validità del parametro stesso. Un altro esempio, seppur meno marcato, della presenza di incertezza è riscontrabile nel parametro  $SCF$  che contiene in sé un'indicazione dell'errore di misura della precipitazione nel caso questa sia neve.

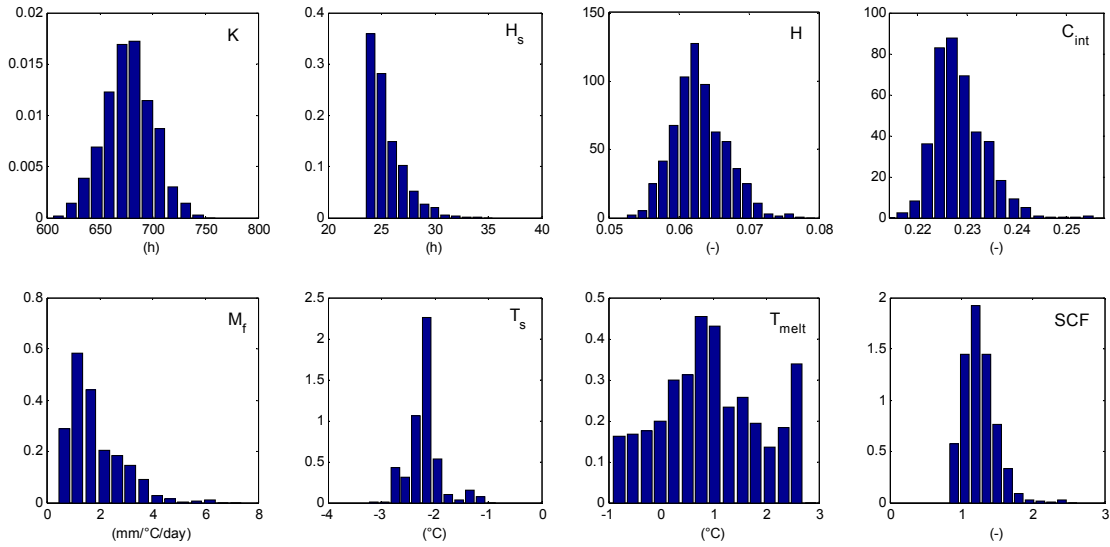


**Figura I- 9:** Risultato della calibrazione del modello AFFDEF sul bacino del torrente Leo presentati in forma di frequenza assoluta e rapportati all'intervallo di variabilità del parametro scelto a priori.

Il parametro invece che è responsabile della distinzione tra le precipitazioni liquide e quelle sotto forma di neve,  $T_s$ , presenta un picco marcato in corrispondenza di  $-2^\circ C$  ed una dispersione dei valori decisamente meno accentuata attorno al picco stesso, significando dunque che non risulta affetto dallo stesso grado di incertezza propria dei parametri  $T_{melt}$  e  $SCF$ ; in ultimo la distribuzione dei valori assunti da  $M_f$ , il fattore gradi-giorno, si condensa in una porzione limitata dell'intervallo di validità, indice probabilmente anche di una corretta scelta dell'intervallo stesso, ed ha una distribuzione ben definita.

Si è soliti in realtà presentare i risultati di queste simulazioni facendo riferimento alle distribuzioni di densità di probabilità a posteriori proprie di ciascun parametro; in Figura I- 10 si trovano i risultati riferiti al torrente Leo. Si noti che in questo caso non solo si è preferito ricorrere ad unità di misura più usuali, ma che non si è deciso di riportare sull'asse delle ascisse l'intero campo di variabilità dei parametri. In questo modo si possono visualizzare meglio le distribuzioni dei valori dei primi quattro parametri, che con la rappresentazione precedente risultavano estremamente penalizzati. Si noti in particolar modo che gli istogrammi riferiti ai parametri  $K$ ,  $H$ ,  $C_{int}$  ricordano le distribuzioni Gaussiane, mentre l'istogramma dei parametri  $H_s$  e  $M_f$  sottolinea alte probabilità per valori prossimi al limite inferiore. Per risolvere i problemi legati ai parametri  $T_{melt}$  e  $SCF$ , immaginando di potere ottenere una distribuzione ben definita anche in questi casi, si potrebbe pensare di aumentare la grandezza dell'intervallo di validità all'interno del quale procedere alla ricerca di soluzioni. In questo modo però si

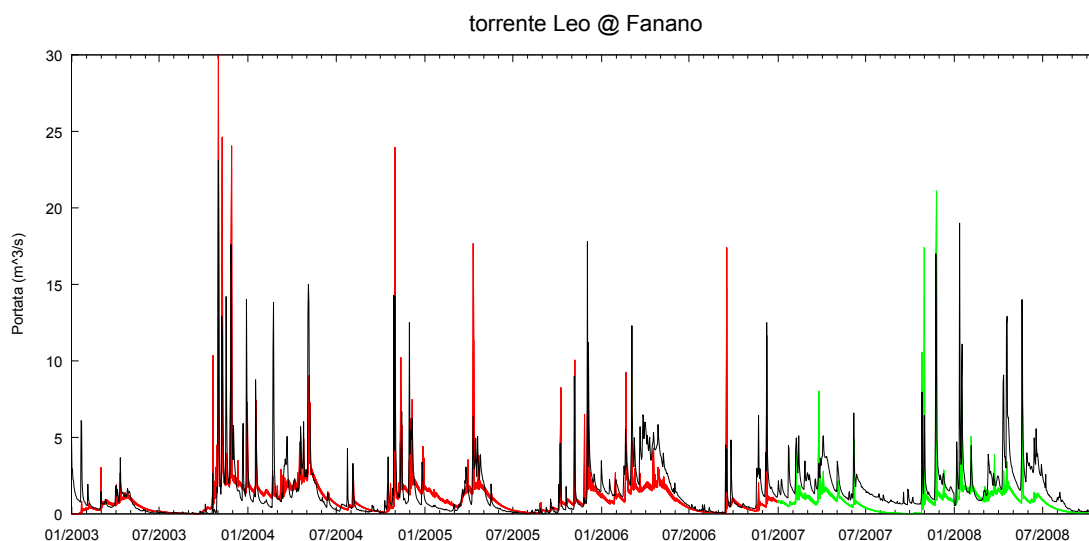
correrebbe il rischio di adottare valori che portano a un comportamento del modello fisicamente non realistico.



**Figura I- 10:** Funzioni densità di probabilità a posteriori dei parametri di AFFDEF calibrato sul bacino del torrente Leo.

Per comprendere quale sia l'effetto dell'incertezza parametrica sulle previsioni di portata si faccia riferimento alla Figura I- 11 nella quale sono presentati i risultati delle simulazioni idrologiche effettuate, in termini di portata simulata. Si determini, per ciascun set di parametri ottenuto dalla calibrazione, la risposta  $S_k = f(\theta_k, \hat{\zeta}, \hat{\phi})$  del bacino sollecitato sempre dalle medesime condizioni meteorologiche (con  $k = 1, \dots, J$  essendo  $J$  il numero di realizzazioni complessive); i risultati delle  $J$  diverse simulazioni possono essere presentati ricorrendo a idrogrammi di portata simulata che vengono espressi sotto forma di fasce di incertezza pari al 90% . Le portate simulate durante il periodo individuato per la calibrazione del modello sono raffigurate in rosso, mentre le simulazioni che si riferiscono al periodo di verifica dell'abilità predittiva del modello sono presentate in verde. I valori della serie di portate osservate sono invece riportati a parte in nero. Si ricordi che le zone dell'idrogramma in cui il dato osservato è compreso all'interno delle fasce di incertezza del simulato sono effettivamente zone nelle quali il modello riesce a ben rappresentare il comportamento del bacino, mentre i problemi sorgono là dove la curva delle portate osservate non interessa la banda simulata. Si vede dunque come il modello non sia in grado di ben rappresentare il comportamento di alcune porzioni dell'idrogramma.

Nella presente trattazione però si è deciso di puntare l'attenzione unicamente sulla stima dell'incertezza parametrica, dunque è bene sottolineare che i risultati ottenuti non tengono in considerazione il ruolo dell'incertezza strutturale del modello, così come tralasciano la presenza di errori nei dati meteorologici.



**Figura I- 11:** Simulazioni idrologiche condotte sul bacino del torrente Leo: valori di portate osservate (in nero) e fasce di incertezza al 90% delle portate simulate. Si distingue il periodo di calibrazione (in rosso) da quello di validazione (in verde)

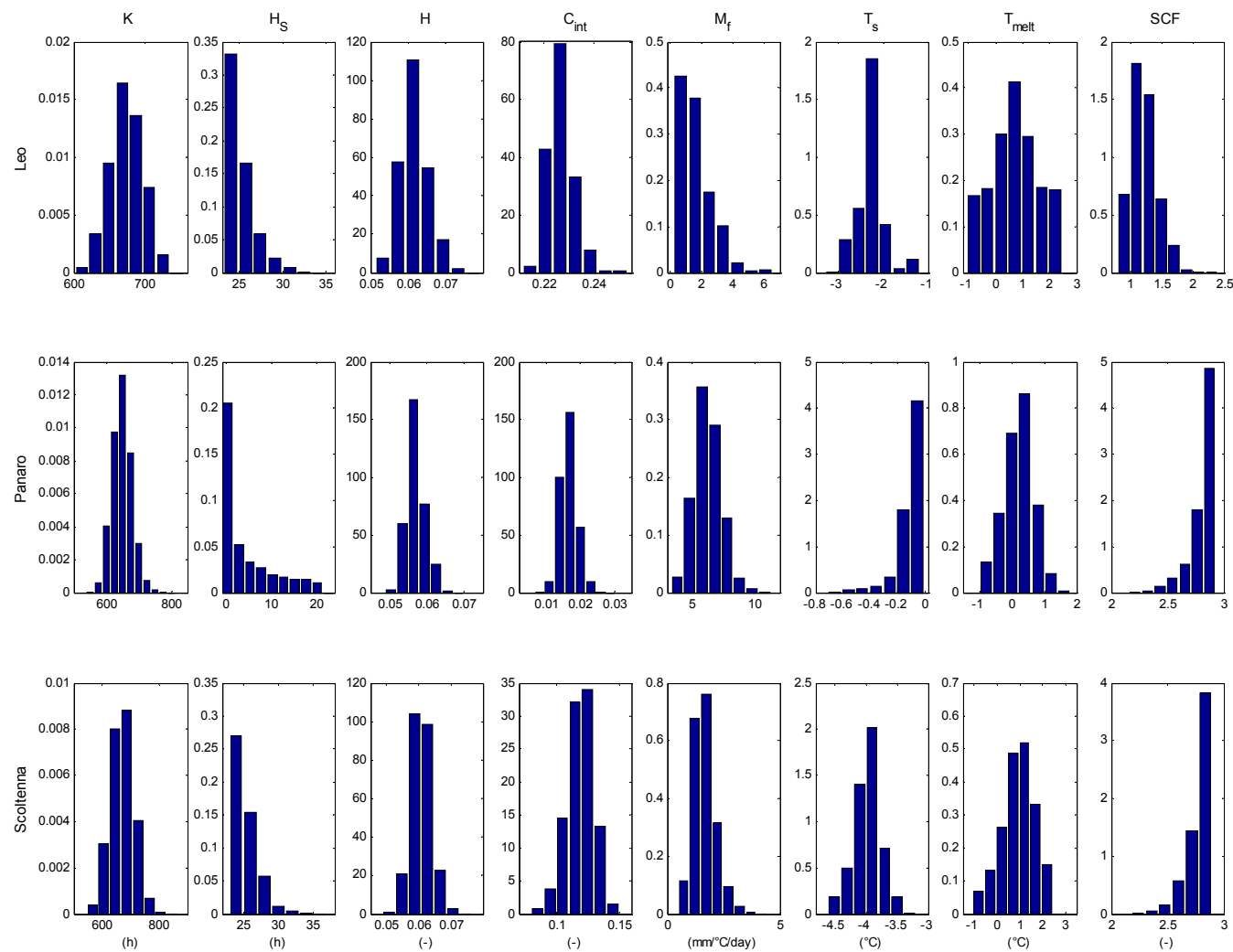
### 3.5 Discussione dei risultati

Con la presente analisi si vuole verificare se sia possibile che alcuni dei parametri del modello afflussi-deflussi scelto possano essere considerati omogenei su svariati bacini; per questo è bene dunque paragonare i risultati ottenuti in termini di distribuzioni di densità di probabilità a posteriori dei parametri sulle diverse aree. In Figura I- 12 sono riportate le distribuzioni di probabilità dei valori assunti da ciascuno degli 8 parametri che caratterizzano il modello AFFDEF, calibrato sul bacino del torrente Leo, del fiume Panaro e del torrente Scoltenna.

Da una prima analisi visiva si vede come i risultati siano promettenti; se si paragonano ad esempio le distribuzioni assunte dal parametro  $K$ , che regola i deflussi che dal serbatoio che rappresenta la falda profonda vengono convogliati direttamente alla sezione di chiusura del bacino, si nota che si tratta in tutti e tre i casi di densità dotate di una forma che ricorda una distribuzione Gaussiana, che assume valori su un intervallo sempre compreso tra le 600 e le 800 ore. Si può dunque immaginare che il tempo necessario per il trasferimento dei deflussi di falda alla sezione di chiusura possa essere considerato sensibilmente costante su tutti i tre bacini considerati e che dunque il parametro  $K$  abbia carattere di omogeneità nello spazio.

Un ragionamento pressoché analogo può essere ripetuto anche per il parametro  $H$ , che, moltiplicato per il valore CN rappresenta la capacità del serbatoio di infiltrazione, diversa non solo di bacino in bacino, ma che, all'interno dello stesso bacino, assume valori differenti in ogni cella di discretizzazione del bacino. Questo dunque non porta a concludere che tutte le zone considerate sono caratterizzate dalla medesima capacità di infiltrazione, ma che la variabilità del fenomeno è riassunta dalla variabilità del valore CN.





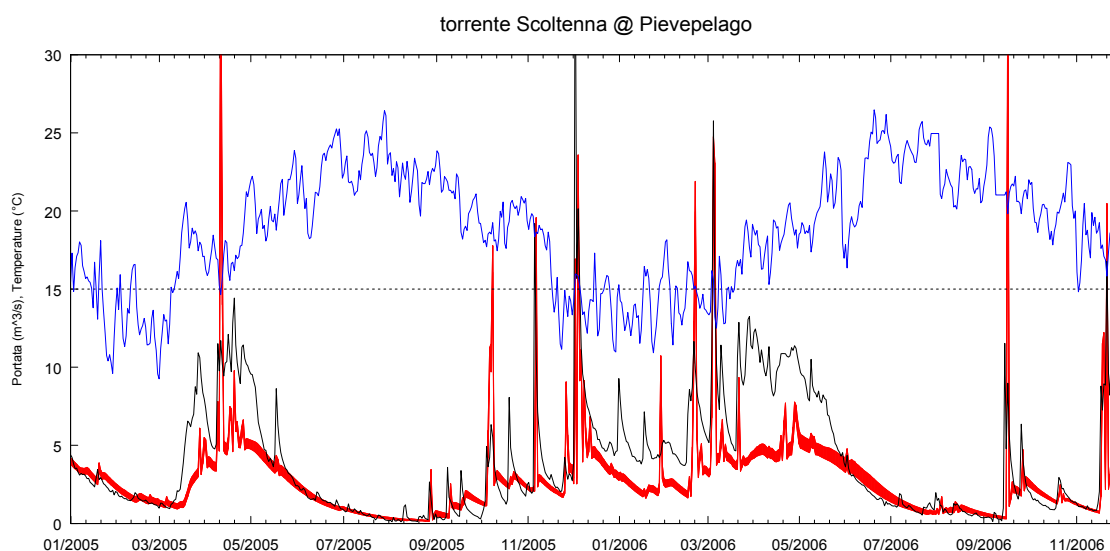
**Figura I- 12:** Funzioni di densità di probabilità a posteriori riferite ai parametri del modello AFFDEF calibrato sui bacini dei torrenti Leo, Scoltenna e del fiume Panaro.

I tempi necessari per il trasferimento dei deflussi verso la falda profonda, descritti dal parametro  $H_s$ , così come la quantità di precipitazione che viene intercettata dalla vegetazione presente in ciascuna cella in cui si è deciso di suddividere i bacini, ed immagazzinata all'interno di un serbatoio la cui capacità massima è rappresentata da  $C_{int} \cdot S(i, j)$ , presentano caratteri simili sulle tre aree analizzate. Il parametro  $H_s$  assume valori con una probabilità maggiore verso il limite inferiore dell'intervallo di validità definito a priori e si ripropone con una distribuzione simile sulle tre aree; si nota però che il bacino del fiume Panaro è caratterizzato da tempi di trasferimento leggermente inferiori (i valori più probabili si condensano tra le 0 e le 5 ore). La distribuzione dei valori che assume il parametro  $C_{int}$  si ripropone sostanzialmente con la medesima forma sui tre bacini considerati, ma presenta valori caratteristici bassi per il bacino del Panaro considerato nel suo insieme; nel caso invece del bacino del torrente Leo, si noti come i valori assunti siano simili allo 0.20 che già il metodo del Curve Number, originariamente proposto dal Soil Conservation service (1972), suggeriva.

Si è già anticipato che i parametri legati al fenomeno di accumulo e scioglimento della neve sono quelli caratterizzati dal maggior grado di incertezza e ne si ha una riprova anche dal confronto delle distribuzioni dei vari parametri su tutti i bacini considerati. In particolare, riferendosi al caso del torrente Leo, si era notato come i due parametri  $SCF$  e  $T_{melt}$  siano quelli ai quali si associa la maggiore incertezza. Lo stesso fatto si ripropone anche sugli altri due bacini considerati; inoltre si noti come nel caso del fiume Panaro e del torrente Scoltenna i valori di  $SCF$  tendano a concentrarsi verso il limite superiore dell'intervallo di validità, dando luogo così a una distribuzione di tipo completamente diverso rispetto a quella che si verifica nel caso del torrente Leo. Per quel che riguarda il parametro  $T_{melt}$  si può notare che in tutti i casi tende ad assumere valori su tutto l'intervallo di validità e presenta un picco centrale.

Il fattore gradi giorno,  $M_f$ , così come la temperatura  $T_s$  con la quale si è scelto di differenziare le precipitazioni liquide da quelle a carattere nevoso, presentano distribuzioni dei valori di tipo differente a seconda del bacino considerato, ed assumono valori in un intervallo ristretto rispetto all'intero campo di validità definito a priori, indice questo di un minor grado di incertezza, rispetto a quella associata ai parametri  $SCF$  e  $T_{melt}$ . Incertezze forti nei parametri che afferiscono al comparto neve sono fisicamente giustificabili; si consideri infatti una tipica serie di osservazioni di temperatura media giornaliera registrate da una stazione di misura presente in questa zona che esemplifica il comportamento tipico delle temperature nell'Appennino emiliano (si veda Figura I- 13). A differenza di quanto ad esempio accade sulle Alpi, nel tardo inverno ed all'inizio della primavera le temperature medie giornaliere cominciano a crescere ma presentano per un periodo lungo una frequente oscillazione attorno al valore di  $0^{\circ}\text{C}$ ; in questo modo dunque la rappresentazione dello scioglimento nivale previsto dal modello deve adeguarsi a continui cambiamenti di stato. Si ricordi poi che ci si sta riferendo a valori medi giornalieri per approssimare un fenomeno che è altamente dipendente da qualsiasi fluttuazione di temperatura avvenga anche nel corso della stessa giornata. Le approssimazioni che si stanno facendo sono dunque molto forti. Sulle Alpi invece queste oscillazioni di temperatura media giornaliera attorno agli  $0^{\circ}\text{C}$  sono molto meno frequenti e dunque il processo di scioglimento nivale risulta decisamente più semplice da simulare.

E' lecito chiedersi anche se le differenze che ci sono tra le distribuzioni che i parametri assumono su differenti bacini riflettano i differenti comportamenti propri di ciascuna area. Si sono già brevemente riportate argomentazioni riguardo i due parametri  $K$  ed  $H$ ; la propensione del fiume Panaro ad assumere valori di  $H_s$  e di  $C_{int}$  inferiori rispetto a quelli caratteristici degli altri due bacini sta a sottintendere che il trasferimento di acqua dal serbatoio di infiltrazione a quello profondo avviene con una maggiore velocità e che la massima quantità di acqua che può essere trattenuta per intercettazione dalla vegetazione è inferiore rispetto a quella che viene trattenuta negli altri due casi. Questi due fatti possono essere spiegati pensando che il bacino del fiume Panaro risulta non solo di dimensioni molto maggiori rispetto a quelle dei bacini dei torrenti Scoltenna e Leo, ma si tratta di un bacino che li comprende al suo interno. E' dunque un'area questa caratterizzata da una variabilità dei processi superiore rispetto alle altre e che, per esempio presenta al suo interno sia zone di alta montagna, caratterizzate da rocce, nelle quali prevalgono i fenomeni di ruscellamento, sia zone boschive sia zone deputate a coltivazioni agricole (nella parte più prossima alla sezione di chiusura). La differente tendenza a consentire il deflusso in ciascuna di queste zone viene rappresentata attraverso la variabilità del parametro CN che assume valori differenti in ciascuna cella  $(i, j)$  che si combina anche a differenti tempi di rilascio di portata verso il serbatoio "profondo" che rappresenta la falda.



**Figura I- 13:** registrazioni di temperatura media giornaliera della stazione di Doccia Fiumalbo (in blu), riferite allo 0 termometrico (linea nera tratteggiata) e valori di portata osservata (in nero) e simulata (in rosso) alla stazione di Pievepelago

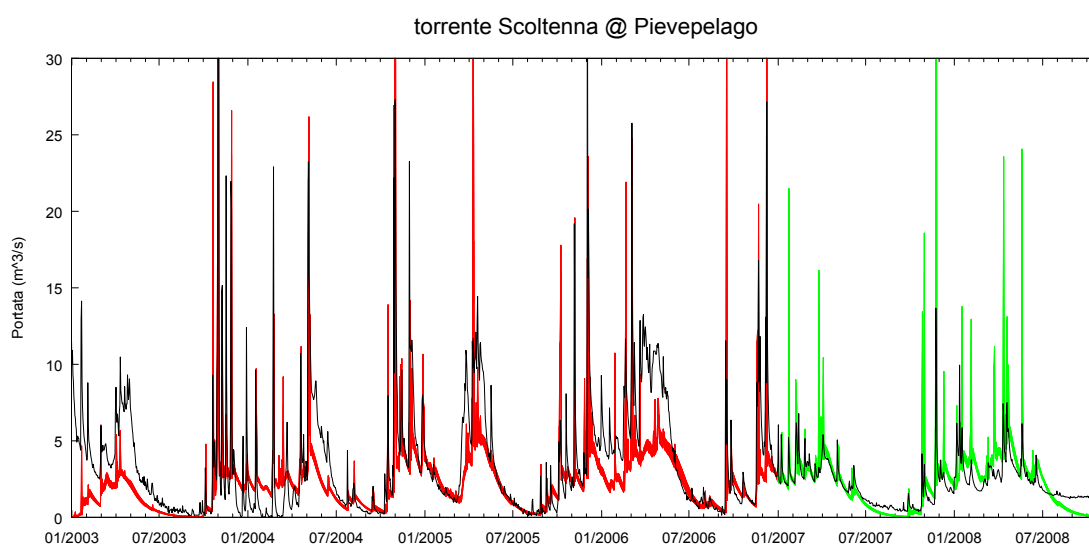
Le considerazioni che riguardano invece il comparto neve derivano sostanzialmente dal fatto che la porzione di territorio che si trova a ridosso del crinale appenninico è quella nella quale i fenomeni di accumulo e scioglimento nivale si presentano in modo più marcato. I bacini del Leo e dello Scoltenna sono dunque direttamente interessati da questi processi. Inoltre, i pluviometri dei quali si dispone non sono tutti riscaldati e gli errori di misura della precipitazione nevosa sono di grande entità. Tutto ciò sta a sottolineare il fatto che i valori bassi di  $SCF$  caratteristici del bacino del torrente Leo possono in parte essere spiegati dal fatto che probabilmente in quell'area si dispone di una strumentazione che riesce a misurare correttamente la quantità di neve che cade al suolo, mentre i valori molto alti del medesimo parametro che si riferiscono al

bacino del torrente Scoltenna, presuppongono la presenza di alti errori di misurazione. Inoltre i valori alti di  $SCF$  registrati siano sul bacino del fiume Panaro considerato nel suo insieme sono legati all'alta variabilità dell'intensità e della frequenza dei fenomeni; solo le porzioni racchiuse all'interno dei bacini del torrente Leo e dello Scoltenna infatti sono abitualmente interessate da ingenti precipitazioni nevose, mentre solo saltuariamente si hanno accumuli di neve al suolo nelle zone a quota inferiore. L'alta correlazione tra tutti i parametri che appartengono al modulo neve può dunque spiegare anche il motivo per cui, sempre il bacino del Panaro sia caratterizzato da valori di  $M_f$ ,  $T_s$  e  $SCF$  che sostanzialmente si discostano da quelli caratteristici dei suoi sottobacini.

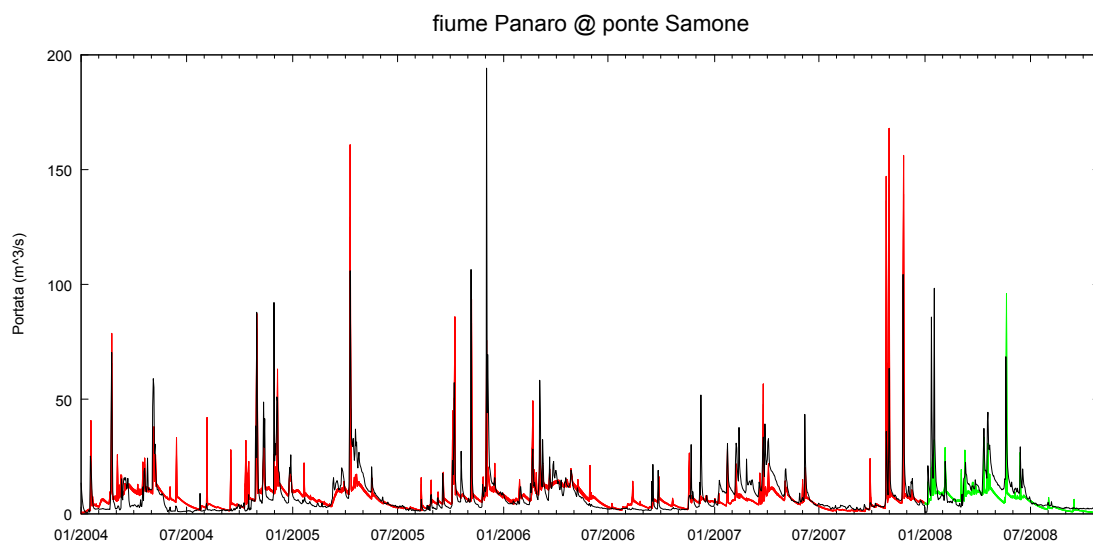
Il ricorso alla suddivisione dell'area in differenti fasce altimetriche all'interno delle quali adottare un comparto neve caratterizzato da parametri di fascia in fascia diversi sarebbe una soluzione per poter superare i problemi di incertezza elevata legati ai parametri coinvolti in questo modulo. Si tratta di una soluzione adottata di sovente (*Nester et al., 2011*) proprio per simulare al meglio i fenomeni che avvengono in zone caratterizzate da una forte escursione altimetrica. Si tratta però anche di una procedura svantaggiosa per l'obiettivo che si è posti con questa trattazione, nel senso che adottando una suddivisione in fasce altimetriche si vanno a moltiplicare anche i parametri che necessitano di essere calibrati.

Si sta infatti realizzando un'analisi di incertezza parametrica al fine di identificare alcuni parametri che, presentando caratteristiche di omogeneità spaziale, non necessitano di essere calibrati su tutti i bacini oggetto di studio, potendo così ridurre il numero complessivo di parametri. Con la suddivisione in fasce altimetriche si andrebbe invece nella direzione opposta.

L'analisi poi degli idrogrammi di portata simulati e osservati riferiti ai casi dei bacini del torrente Scoltenna e del fiume Panaro ripropongono le problematiche viste nel caso del bacino del torrente Leo: alcune parti dell'idrogramma non possono essere ben rappresentate dal modello adottato. Ci si riferisca alla Figura I- 14 e alla Figura I- 15.



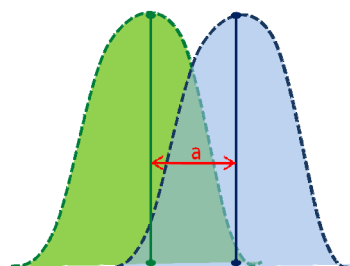
**Figura I- 14:** Simulazioni idrologiche condotte sul bacino del torrente Scoltenna chiuso alla sezione di Pievepelago



**Figura I- 15:** Simulazioni idrologiche condotte sul bacino del fiume Panaro chiuso alla sezione di ponte Samone

### 3.6 Analisi dell'ipotesi di invarianza spaziale dei parametri

Con l'analisi di incertezza parametrica proposta si sono stimate le distribuzioni di densità di probabilità dei parametri; queste risultano fisicamente verosimili e sono ben definite.



**Figura I- 16:** distribuzioni di densità di probabilità di un parametro che ipoteticamente è stato calibrato su due differenti bacini

Da una prima analisi visiva si è dunque visto che le distribuzioni di probabilità di quasi tutti i parametri, eccezion fatta per  $\alpha$ ,  $\beta$ ,  $\gamma$ , sono in diversa misura compatibili. In taluni casi le distribuzioni appaiono compatibili sui tre bacini; è ad esempio quello che accade alle distribuzioni di densità di probabilità a posteriori che si riferiscono al parametro  $\alpha$  calibrato sui tre bacini. Da un punto di vista statistico la compatibilità dei parametri è quantificata dalla area comune alle rispettive distribuzioni di probabilità (come esemplificato in Figura I- 16). Infatti, la misura dell'area di sovrapposizione fornisce un indice variabile fra 0 e 1, limiti ai quali corrispondono, rispettivamente, l'assenza e la totale compatibilità.

Poiché la compatibilità, come peraltro ci si attendeva, non è in nessun caso totale, ne discende che l'introduzione dell'ipotesi di invarianza spaziale dei parametri introduce a sua volta un'approssimazione nei risultati che è necessario quantificare. Ciò è possibile imponendo che i parametri assumano effettivamente gli stessi valori su bacini diversi e successivamente

verificando le prestazioni dei modelli. Infatti, nelle analisi precedenti, il modello ha potuto adattarsi di volta in volta, a seconda del bacino considerato, alle particolari condizioni idrologiche dell'area in esame. Era quindi prevedibile che i set di parametri che portano alle soluzioni più verosimili si trovassero in aree diverse dello spazio dei parametri a seconda del caso studiato, ma lo studio eseguito non fornisce indicazioni in merito alla perdita di prestazione che si ottiene imponendo valori costanti dei parametri stessi.

In realtà, esistono situazioni pratiche nelle quali è effettivamente necessario assumere valori dei parametri invariati su diversi bacini. Ad esempio, nel caso di bacini non strumentati, per i quali è impossibile procedere ad una calibrazione locale del modello e si ricorre dunque all'adozione di parametri che derivano dalla calibrazione su un bacino idrologicamente simile. Quali sarebbero le conseguenze di una simile ipotesi è il tema che verrà affrontato nel prossimo capitolo.

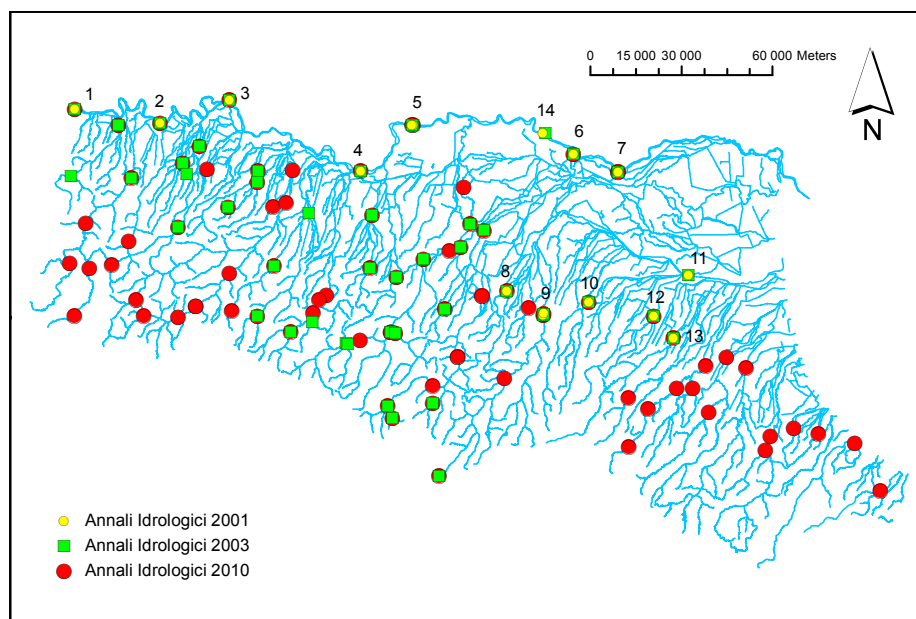
## **4 Stima dell'approssimazione indotta dall'adozione di parametri omogenei nello spazio.**

Sin qui si è infatti parlato unicamente di calibrazione, ma si vuole meglio comprendere che cosa succede nel momento in cui si passa alla fase di previsione, cioè quali siano i benefici, in fase di validazione, derivanti dall'utilizzo di un set di parametri invariato contemporaneamente su differenti bacini. Si ha la convinzione che, soprattutto in caso di “data scarcity”, la robustezza di un modello aumenti all'aumentare delle informazioni che vengono adottate nella taratura dei parametri del modello stesso; è anche questo il motivo per cui in calibrazione si tende ad utilizzare un periodo sensibilmente lungo. Così il modello afflussi-deflussi viene sollecitato da varie e differenti condizioni meteorologiche, in conseguenza delle quali si verificano sia magre, che piene repentine, che morbide dovute ad esempio allo scioglimento nivale; sono tutte condizioni alle quali il modello si adatta ed i suoi parametri vengono aggiustati in modo che, in fase di previsione, abbia memoria di tutti i fenomeni e li sappia riconoscere. Si vuole sperimentare se, incrementando il numero e il tipo di informazioni meteorologiche adottate nella calibrazione di un modello, sia possibile sperimentare i benefici di questa scelta in fase di previsione. O, più esplicitamente, se si sceglie di adottare lo stesso modello, uguale non solo per struttura, ma utilizzando anche gli stessi parametri, su un'area più estesa rispetto a quella individuata da un unico bacino, è possibile sollecitare e tarare il modello con informazioni meteorologiche che, provenendo da aree differenti, sono di diverso tipo. Il modello così calibrato, se applicato ad un unico bacino, avrà, probabilmente, performance peggiori in fase di calibrazione rispetto a quanto accadrebbe se venisse tarato adottando unicamente le informazioni provenienti dalla zona identificata dal bacino di interesse; questo fatto è giustificabile sottolineando che la calibrazione effettuata non ha l'obiettivo di adattare il modello solo ed unicamente alle condizioni che si verificano sul bacino scelto. In validazione, quando il modello viene applicato al bacino di interesse, il modello dovrà adattarsi a condizioni meteorologiche nuove che, probabilmente, ha già sperimentato sugli altri bacini coinvolti nella fase di calibrazione.

Esiste poi una seconda motivazione che ha spinto alla verifica delle conseguenze derivanti dalla adozione di parametri costanti su un'area di dimensioni più estese rispetto a quelle di un unico bacino: il caso di bacini non strumentati.

## 4.1 Bacini non strumentati

Sivapalan et al. (2003) definiscono un bacino non strumentato, come “un bacino dotato di serie temporali di misure non adeguate (sia dal punto di vista quantitativo che qualitativo) delle variabili idrologiche, così da rendere impossibile la stima delle variabili idrologiche di interesse (che rappresentino sia la qualità che la quantità di acqua ivi presente) all’appropriata scala temporale e spaziale anche mediante l’uso di modelli matematici”; gli stessi autori continuano poi sottolineando il concetto con un esempio: “se la variabile di interesse non è stata misurata con la risoluzione spaziale richiesta, o durante il periodo di tempo richiesto per effettuare previsioni o per la calibrazione del modello, il bacino viene allora classificato come non strumentato nei riguardi di quella stessa variabile”. I bacini che ricadono in questa definizione sono molti in tutto il mondo ed anche in regione Emilia Romagna la situazione risulta critica.



**Figura I- 17:** Consistenza della rete delle stazioni di misura di portata in Emilia Romagna al 31/12/2001 (in giallo), 31/12/2003 (in verde), 31/12/2010 (in rosso); (R.E.R., 2001, 2003, 2010).

A riprova del fatto che la rete di monitoraggio dei fiumi della regione Emilia Romagna sia particolarmente “giovane”, si noti come al 31/12/2001 le uniche stazioni di misura attive fossero quelle riportate in giallo in Figura I- 17; il fiume Po era dotato di otto stazioni di misura che, peraltro, registrano le osservazioni di portata da lunga data, come riportato in Tabella I- 4.

	Stazione	Inizio osservazioni (anno)
1	Po a Spessa	1991
2	Po a Piacenza	1869
3	Po a Cremona	1870
4	Po a Boretto	1943
5	Po a Borgoforte	1923
6	Po a Ficarolo	1988
7	Po a Pontelagoscuro	1807
8	Samoggia a Calcara	1937
9	Reno a Casalecchio	1851
10	Idice a Castenaso	1975
11	Reno a Bastia	1929
12	Sillaro a Sesto Imolese	1986
13	Santerno a Mordano	1989
14	Po a Sermide	1880

**Tabella I- 4:** Elenco delle stazioni di misura di portata in funzione al 31/12/2001. (R.E.R., 2001)



Solo a partire dal 2003, come testimonia l'annale idrologico di quell'anno, la rete degli strumenti di misura di portata prevede stazioni anche sugli affluenti emiliani del fiume Po, che cominciano dunque ad essere monitorati. Un ultimo sguardo all'annale idrologico del 2010 mostra invece come, nel corso degli anni, la rete di strumentazione sia stata potenziata ed anche i fiumi della Romagna siano stati dotati di stazioni di misura. Si può dunque vedere come il problema dei bacini non strumentati sia effettivamente di grande attualità anche in questa regione.

Qualsiasi modello idrologico necessita, in qualche modo, di una procedura di calibrazione dei parametri per poter produrre simulazioni affidabili; come si è mostrato nei capitoli precedenti, si tratta di un processo durante il quale i parametri del modello scelto vengono tarati fino a quando la serie di portate simulate riproduce al meglio quelle osservate (*Todini, 1988*). In assenza di misure di portata, registrate su un arco temporale sufficientemente lungo e coincidente con quello sul quale sono stati registrati i dati meteorologici che vengono utilizzati come ingresso del modello afflussi-deflussi, risulta quindi impossibile effettuare la procedura di calibrazione stessa. Nel caso dunque la sezione di interesse sia scarsamente strumentata o non strumentata è necessario identificare i valori assunti dai parametri del modello, ricorrendo a metodi alternativi (*Montanari & Toth, 2007*). Solitamente si può scegliere di affrontare il problema seguendo due approcci totalmente differenti: si può ricorrere all'uso di modelli idrologici di tipo fisicamente basato oppure si può procedere alla regionalizzazione dei parametri del modello sfruttando le caratteristiche fisiche del bacino. La scelta di utilizzare modelli fisicamente basati è giustificata dal fatto che i parametri adottati possono essere strettamente legati a proprietà fisiche del bacino e risultano dunque misurabili. Nella pratica, però, a causa dell'eccessivo numero di parametri che solitamente caratterizza questi modelli e, dunque, degli inevitabili errori strutturali del modello stesso, è comunque necessario procedere alla calibrazione di alcuni parametri (*Yadav et al., 2007*).

Si propone dunque come metodo alternativo alla calibrazione regionale, l'adozione di parametri che provengano semplicemente dalla calibrazione del modello su altri bacini che siano però strumentati. Si procede poi al trasferimento dei parametri stessi direttamente al caso non strumentato, senza effettuare ulteriori trasformazioni.

## **4.2 L'ipotesi di omogeneità dei parametri: la procedura di calibrazione**

Per verificare l'approssimazione indotta dall'adozione dell'ipotesi di omogeneità dei parametri si ricorre allo studio della medesima area e dello stesso modello afflussi-deflussi scelto per l'analisi dell'incertezza parametrica proposta al capitolo precedente. Si ricorre dunque all'applicazione del modello AFFDEF sull'area composta dai bacini dei torrenti Leo e Scoltenna (chiusi rispettivamente nelle sezioni di Fanano e Pievepelago) e del fiume Panaro sotteso alla sezione di ponte Samone (si veda Figura I- 6). Si dispone, come in precedenza, di dati di portata misurati a passo giornaliero tra il 01/01/2003 e il 31/12/2008 in corrispondenza delle tre sezioni fluviali oltre che di registrazioni di temperatura media giornaliera e di pioggia giornaliera cumulata.

A differenza della metodologia precedente, in questo caso si rivolge l'interesse alla ricerca dell'insieme di parametri "ottimale", quel set cioè che rende il modello in grado di meglio simulare il comportamento del bacino in esame, trascurando in questa occasione l'influenza di tutte le incertezze in gioco (incertezza strutturale del modello, incertezza parametrica e incertezza nei dati in ingresso).

L'analisi che si va ora a proporre ha un duplice obiettivo; da un lato si vuole verificare quali siano le conseguenze, sia in calibrazione che in validazione dell'adozione di parametri costanti su tutta l'area di studio. D'altro canto, immaginando di voler simulare il comportamento di un bacino non strumentato, si vogliono verificare le prestazioni di un modello che adotta gli stessi parametri calibrati su un bacino differente, a lui prossimo e strumentato.

Si è deciso di sviluppare l'analisi secondo i seguenti passi:

- a. Si adotta come punto di riferimento il risultato della calibrazione (e della validazione) del modello AFFDEF su ciascun bacino, considerato a sé stante, per il quale si determina l'insieme di parametri che meglio si adatta alle condizioni meteorologiche e idrologiche tipiche del bacino. Si effettua cioè una classica calibrazione dei parametri.
- b. Per svincolarsi dalla scelta di un particolare periodo di calibrazione (e dunque di validazione), si ripete l'analisi svariate volte; tutte le simulazioni effettuate su ciascun bacino si differenziano per la scelta del periodo di validazione (1 anno solare) che cambia di volta in volta.
- c. Si assume a questo punto l'ipotesi di omogeneità dei parametri su una regione di grandezza superiore a quella coperta da un singolo bacino idrografico e si verificano le conseguenze di questa scelta man a mano che aumentano le dimensioni della regione di omogeneità:
  - in prima battuta si considera una regione composta unicamente dai due bacini dei torrenti Leo e Scoltenna;
  - il secondo passaggio prevede l'utilizzo di tutta l'area costituita dai bacini dei torrenti Leo, Scoltenna e del fiume Panaro.
  - Qualunque sia la grandezza dell'area esaminata, si procede alla calibrazione del modello, contemporaneamente e considerando invariati i valori di tutti i parametri del modello, su tutti i bacini che ricadono nell'area. Così come era accaduto per il caso della calibrazione di ogni bacino considerato a sé stante, anche in questo caso si sceglie di svincolarsi dalla scelta del periodo di validazione e si ripete svariate volte questo processo considerando ad ogni passaggio un anno differente ai fini della validazione.
- d. Si paragonano ora i risultati ottenuti dalla classica calibrazione (e validazione) dei bacini e quelli delle analisi realizzate considerando i parametri costanti su più bacini contemporaneamente.

Da ultimo poi, si considera il caso dei bacini non strumentati; si ipotizza cioè che un bacino a turno non disponga di misure di portata nella sezione di chiusura e che dunque sia impossibile procedere alla calibrazione del modello in quella sezione. Si verificano quindi le prestazioni del modello che adotta i medesimi valori dei parametri che sono stati calibrati su un bacino a lui prossimo.

Per completare la descrizione della particolare procedura di calibrazione adottata rimangono da definire alcune scelte fondamentali:

- ✓ Qualsiasi simulazione si stia effettuando, si adotta un periodo di riscaldamento per il modello sempre pari a 300 giorni, così da trascurare gli effetti legati alla scelta delle condizioni iniziali
- ✓ La funzione obiettivo che deve essere massimizzata è rappresentata dal valore dell'efficienza calcolata secondo Nash (*Nash & Sutcliffe, 1970*):

$$NSEff = 1 - \frac{\sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \overline{Q_{obs}})^2} \quad (4.1)$$

nella quale  $Q_{obs,i}$ ,  $Q_{sim,i}$  e  $\overline{Q_{obs}}$  sono rispettivamente i valori di portata osservata e simulata al tempo  $i$  e la media delle portate osservate. L'efficienza di Nash così calcolata può assumere valori tra 1 e  $-\infty$ ; in particolare si possono verificare alcune condizioni tipiche:

- $NSEff = 1$  si verifica quando il modello riesce a riprodurre perfettamente le serie di dati osservati,
- $NSEff = 0$  si presenta quando il modello proposto raggiunge performance che sono uguali a quelle che si otterrebbero se si adottasse come “modello” la semplice media dei valori osservati,
- $NSEff < 0$  si presenta quando il modello afflussi-deflussi non è in grado di simulare a dovere i dati osservati e la media stessa delle osservazioni di portata costituisce un “modello” migliore rispetto a quello adottato.

Nel caso si stia procedendo alla calibrazione simultanea di due (oppure  $n$ ) bacini, la funzione obiettivo da massimizzare risulta invece la media delle efficienze di Nash calcolate sui rispettivi bacini:

$$NSEff = \left( \sum_{i=1}^n NSEff_i \right) / n \quad (4.2)$$

- ✓ L'algoritmo di ottimizzazione, GENOUD (*Mebane et al. 1998*), un algoritmo genetico caratterizzato da tempi di calcolo particolarmente contenuti anche quando si debbano calibrare modelli dotati di numerosi parametri, ed implementato in ambiente R, viene adottato per selezionare l'insieme di valori ottimali dei parametri da utilizzare in ogni caso.
- ✓ Si è scelto inoltre di valutare la bontà del modello utilizzando sempre come indice il valore assunto dall'efficienza calcolata secondo Nash.

### 4.3 Discussione dei risultati

Periodo di Validazione	Fase 1						Fase 2				Fase 3					
	Leo @ Fanano		Scoltenna @ Pievepelago		Panaro @ ponte Samone		Scoltenna @ Pievepelago - Leo @ Fanano				PANARO (ponte Samone-Fanano-Pievepelago)					
							Leo @ Fanano		Scoltenna @ Pievepelago		Leo @ Fanano		Scoltenna @ Pievepelago		Panaro @ ponte Samone	
	C	V	C	V	C	V	C	V	C	V	C	V	C	V	C	V
2003	0.457	0.655	0.514	0.253	ND	ND	0.372	0.676	0.507	0.257	ND	ND	ND	ND	ND	ND
2004	0.545	0.403	0.549	0.432	0.634	0.475	0.518	0.339	0.548	0.432	0.392	0.365	0.525	0.454	0.544	0.406
2005	0.500	0.234	0.525	0.468	0.671	0.441	0.479	0.264	0.487	0.536	0.447	0.355	0.474	0.473	0.613	0.381
2006	0.514	0.518	0.487	0.474	0.627	0.270	0.481	0.456	0.455	0.506	0.479	0.396	0.421	0.407	0.589	0.363
2007	0.557	0.067	0.533	0.022	0.634	0.289	0.510	0.162	0.518	-0.127	0.522	0.139	0.467	0.110	0.593	0.444
2008	0.524	0.411	0.592	-3.556	0.560	0.507	0.489	0.368	0.590	-3.556	0.496	0.398	0.546	-3.517	0.509	0.636
Periodo 2003-2006	Media		0.453		0.407		0.395		0.434		0.433		0.372		0.445	
	Min		0.234		0.253		0.270		0.264		0.257		0.355		0.407	
	Max		0.655		0.474		0.475		0.676		0.536		0.396		0.473	
Periodo 2003-2008	Media tot		0.381		-0.318		0.396		0.377		-0.325		0.331		-0.415	
	Min tot		0.067		-3.556		0.270		0.162		-3.556		0.139		-3.517	
	Max tot		0.655		0.474		0.507		0.676		0.536		0.398		0.473	

**Tabella I- 5:** Valori dell'efficienza di Nash calcolati in calibrazione (C) e in validazione (V) per le simulazioni idrologiche realizzate seguendo l'ipotesi di omogeneità dei parametri nello spazio.

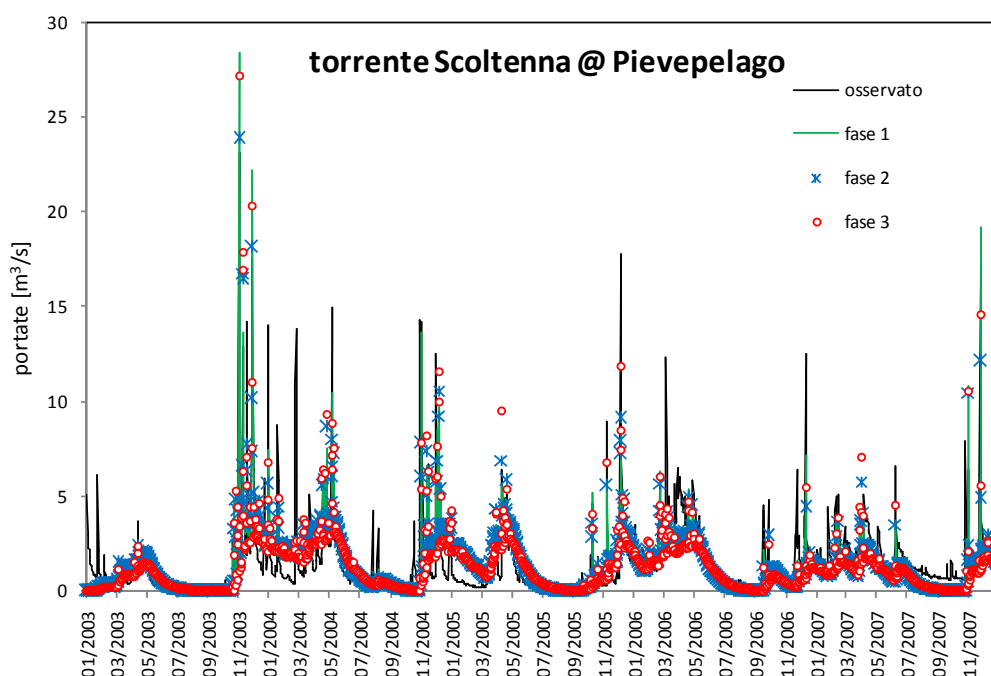
I risultati dell'analisi proposta vengono riportati in Tabella I- 5. Per semplicità si è scelto di schematizzarli riferendosi a tre differenti fasi dell'analisi proposta e definendole:

- “Fase 1”, la prima parte dell'analisi nella quale si calibra (e si valida) il modello su ogni bacino considerato a sé stante;
- “Fase 2”, il successivo passo nel quale si assume valida l'ipotesi di costanza dei parametri sui bacini dei torrenti Leo e Scoltenna che vengono dunque calibrati adottando parametri con gli stessi valori su entrambi i bacini;
- “Fase 3”, l'evoluzione ulteriore nella quale si realizzano simulazioni idrologiche per i bacini dei torrenti Leo e Scoltenna e del fiume Panaro calibrati e validati tutti insieme.

Non disponendo la sezione di ponte Samone di misure di portata osservate durante il 2003, è impossibile effettuare simulazioni sul bacino del Panaro adottando proprio quell'anno come periodo di validazione; per lo stesso motivo non si possono effettuare simulazioni idrologiche sull'area composta da tutti e tre i bacini riferendosi al 2003 (si indica con ND, l'assenza di dati). L'ulteriore conseguenza dell'assenza di dati di portata consiste nell'adottare per il fiume Panaro un periodo di calibrazione sempre pari a 4 anni, a fronte dei 5 anni adottati per i rimanenti bacini.

Si analizzino anzitutto le conseguenze dell'adozione di parametri costanti riferendosi al caso della calibrazione (si seguano dunque, di bacino in bacino, le colonne indicate con C nella tabella); considerando ad esempio il 2004 come periodo di validazione (AFFDEF viene dunque calibrato sugli anni 2003-2005-2006-2007-2008) e seguendo ciò che accade al bacino del Leo, si nota che, di fase in fase, la bontà del modello diminuisce; ciò è una conferma del fatto che all'aumentare delle dimensioni dell'area considerata, e dunque all'aumentare delle informazioni adottate in fase di calibrazione, il modello non riesce a simulare al meglio le condizioni che si verificano unicamente sul bacino oggetto di interesse durante il periodo di calibrazione.

Il comportamento appena descritto per il caso del torrente Leo si ripropone con le medesime caratteristiche anche in tutti gli altri casi. Si riporta come ulteriore esempio il caso del torrente Scoltenna calibrato sul periodo compreso tra il 2003 e il 2007; in Figura I- 18 si vede quali siano, sull'idrogramma di portata riferito unicamente al periodo di calibrazione, gli effetti indotti dalla procedura che prevede le tre fasi sopra citate. Il peggioramento esiste anche se solitamente è modesto e ciò è constatabile sia in termini di bontà del modello che grazie ad un'analisi visiva degli idrogrammi.



**Figura I- 18:** il caso del torrente Scoltenna calibrato, seguendo le tre fasi, sul periodo 2003-2007

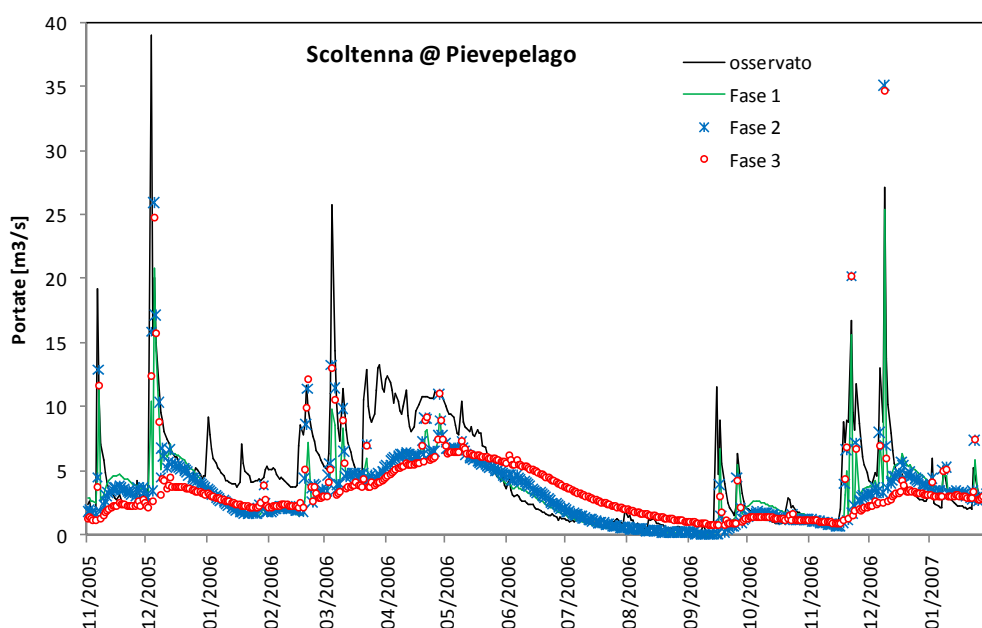
Si può dunque affermare che all'aumentare del numero di bacini coinvolti nella calibrazione si ha una riduzione del valore dell'efficienza, indicando dunque che la bontà della calibrazione

peggiora di fase in fase. Ci si può cioè aspettare che all'aumentare delle dimensioni dell'area di studio, si stia considerando contemporaneamente sia il bacino principale che i suoi sottobacini, il modello perda la flessibilità necessaria per ben adattarsi ai fenomeni che accadono all'interno delle diverse zone.

Non si riesce invece a identificare un comportamento univoco per quanto riguarda la validazione. Si consideri ad esempio il caso del bacino del Leo validato riferendosi all'anno 2004; si passa da un'efficienza di Nash pari a 0.403, ottenuta dalla validazione propria della fase 1, ad efficienze pari a 0.339 e 0.365 caratteristiche della validazione effettuata in fase 2 e fase 3 rispettivamente. Se si considera sempre il bacino del torrente Leo, validato però riferendosi all'anno 2005 si vede come la bontà del modello ottenuto dalla calibrazione contemporanea dei bacini dei torrenti Leo e Scoltenna sia superiore sia al caso di validazione del modello calibrato sia in fase 1 che in fase 3, a riprova del fatto che non è possibile identificare un'unica regola. Se si considerano poi i valori delle efficienze in validazione tipiche degli anni 2007 e 2008 si nota come queste calino drasticamente, assumendo talvolta valori negativi; in questi periodi si sono verificate infatti condizioni idrologiche particolari.

Per comprendere se nel comportamento medio si possa riconoscere un andamento tipico dell'indice utilizzato per descrivere la bontà del modello, si ricorre all'analisi dei valori medi di efficienza calcolati sia sul periodo 2003-2008 che sul periodo 2003-2006, proprio per evitare di considerare i casi estremi che si verificano negli anni 2007 e 2008. Nuovamente si registra un comportamento completamente altalenante; non si riesce dunque ad identificare una regola valida per tutti i casi. Non si può però trarre la conclusione che, all'aumentare del numero di bacini coinvolti nell'analisi il valor medio dell'efficienza calcolata in validazione diminuisca.

In Figura I- 19 si riporta poi un esempio di quali siano le conseguenze dell'adozione dell'ipotesi di omogeneità dei parametri sugli idrogrammi simulati in fase di validazione.



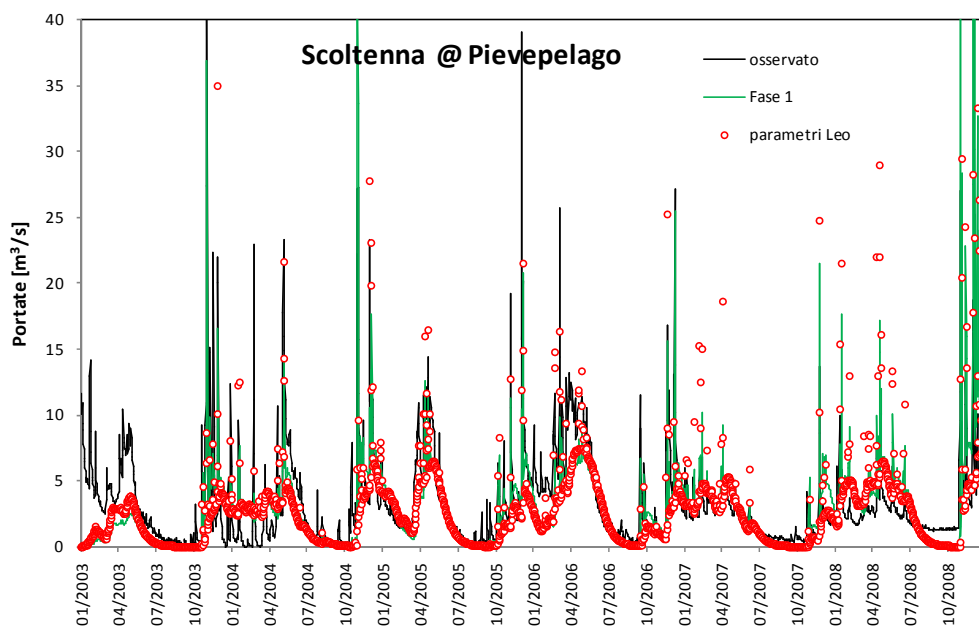
**Figura I- 19:** Il caso del torrente Scoltenna validato sull'anno 2006.

Se si paragonano gli idrogrammi simulati (in validazione) dai modelli messi a punto nelle fasi 1 e 2 non si notano grandi differenze; ciò è probabilmente dovuto al fatto che in fase 2 viene considerato il torrente Leo, il cui bacino ha caratteristiche simili a quelle del bacino dello Scoltenna. Entrambi infatti hanno quota media paragonabile, entrambi comprendono quasi unicamente zone di media – alta montagna, e l’area drenata è sostanzialmente di pari dimensioni. Nel momento in cui anche il bacino del fiume Panaro entra in gioco (nella fase 3), allora il modello produce anzitutto curve di esaurimento di durata maggiore, cioè il serbatoio che rappresenta la falda risulta rallentato.

Si vuole ora capire che cosa accade se si considera un bacino alla volta come se fosse non strumentato. In questo caso si decide di adottare il set ottimale di parametri individuato dalla calibrazione di un altro bacino strumentato ad esso prossimo. Si presenta sempre il caso del bacino del torrente Scoltenna, per il quale si paragonano gli idrogrammi di portata simulata dal modello AFFDEF calibrato sul bacino stesso, ed assunto come condizione base alla quale riferirsi, e quello di portata simulata dal modello ottenuto adottando i parametri che derivano dalla calibrazione del bacino del Leo. In entrambi i casi ci si riferisce al 2006 come periodo di validazione. Si riportano in Tabella I- 6 i valori dei parametri che il modello assume se viene calibrato secondo la Fase 1 e quelli che invece sono stati calibrati sul bacino del torrente Leo e sono stati semplicemente trasferiti sul torrente Scoltenna. Si noti come la stessa tabella riporti anche i valori dei parametri adottati per gli idrogrammi rappresentati in Figura I- 19.

		Fase 1	Fase 2	Fase 3	valori del fiume Leo
K	Days	21.715	24.016	49.385	21.691
Hs	days	6.197	5.890	5.512	5.443
H	[-]	0.477	0.299	0.288	0.187
cint	[-]	0.349	0.498	0.557	0.578
Mf	mm/°C/day	2.592	2.360	2.283	3.982
Ts	°C	-1.163	-1.499	-1.196	-0.758
Tmelt	°C	2.255	1.151	0.411	0.780
SCF	[-]	1.174	1.792	2.229	2.931

**Tabella I- 6:** Valori assunti dai parametri del modello AFFDEF calibrato sul bacino del torrente Scoltenna sul periodo (2003-2005, 2007-2008) secondo le tre diverse fasi previste dall’analisi e valori degli stessi parametri ottenuti dalla calibrazione del modello sul bacino del fiume Leo (allo scopo di verificare il comportamento del bacino dello Scoltenna come se fosse non strumentato).



**Figura I- 20:** torrente Scoltenna (anno di validazione 2006): simulazioni idrologiche effettuate in accordo alla fase 1 e sfruttando il set di valori ottimali ottenuti per il torrente Leo.

Si dimostra così che nel caso di bacini non strumentati, per ottenere una prima simulazione idrologica di buona qualità si può ricorrere all'utilizzo di un set di parametri derivante dalla calibrazione del modello in un bacino strumentato posto nelle vicinanze di quello di interesse.



## 5 Considerazioni conclusive della I parte

La variabilità spaziale dei parametri dei modelli afflussi-deflussi spazialmente distribuiti è un problema ancora poco studiato, soprattutto perché richiede un notevole numero di simulazioni. Nella presente sezione si è affrontato questo tema ricorrendo all'applicazione di AFFDEF, un modello spazialmente distribuito dotato di un contenuto numero di parametri in modo da ridurre lo sforzo computazionale, così da rendere possibile la sua applicazione anche ad una regione di dimensioni vaste.

Si è analizzata l'ipotesi di ridurre il numero dei parametri di un modello afflussi-deflussi che necessitano di essere calibrati cercando di comprendere se sia cioè plausibile che un parametro, ad esempio quello che regola la capacità del serbatoio di infiltrazione previsto dal modello AFFDEF, sia utilizzato su un bacino diverso da quello sul quale è avvenuta la sua calibrazione. Se cioè un parametro risulta non solo invariato nel tempo, ma anche nello spazio, mostrando così carattere di uniformità su una regione ben definita, lo si può escludere dalla procedura di calibrazione che deve essere effettuata su ciascun bacino e la modellazione idrologica di un'area di vaste dimensioni richiederebbe un minor numero di passaggi, rendendo il processo più veloce ed agile. E' dunque questo l'obiettivo che ci si è posti con la presente trattazione.

La stima dell'incertezza parametrica permette di dedurre informazioni riguardo la fattibilità dell'idea di partenza; se le funzioni densità di probabilità a posteriori dei valori assunti da un particolare parametro risultano statisticamente compatibili su differenti bacini, allora si può concludere che il parametro in questione ha un comportamento statisticamente omogeneo sull'area considerata e lo si può dunque escludere dalla calibrazione che deve essere ripetuta per ogni bacino. L'area di studio, costituita dai bacini dei torrenti Leo e Scoltenna e dal bacino del fiume Panaro, nella sua limitatezza presenta le caratteristiche proprie dei bacini appenninici della regione Emilia-Romagna e può essere adottata come area nella quale testare l'idea proposta. L'analisi effettuata porta ad ottenere un primo importante risultato: i valori assunti dai parametri sono caratterizzati da distribuzioni di densità ben definite; i parametri affetti da maggiore incertezza assumono valori che spaziano quasi uniformemente su tutto il dominio di validità e sono proprio quelli che si riferiscono al modulo neve inserito all'interno del modello. Sono soprattutto i parametri  $T_{melt}$  e  $SCF$ , che sanciscono rispettivamente l'inizio della fase di scioglimento nivale e l'errore associato alle misure della precipitazione che cade in forma solida, che risentono di grande incertezza, a riprova del fatto che simulare la presenza di neve in Appennino è decisamente complicato, ancor più, ad esempio, di quanto accade sulle Alpi. Alcuni parametri sono caratterizzati da distribuzioni di densità di probabilità a posteriori che, riferite ai diversi bacini considerati, risultano visivamente molto simili tra loro ed in parte sovrapponibili. La compatibilità tra i parametri risulta in nessun caso totale, come d'altronde ci si attendeva. Fin ad ora si è analizzato il comportamento del modello solo in fase di calibrazione. Si pensi però che esistono situazioni particolari nelle quali proprio non si può

applicare una calibrazione rigorosa e dunque l'unica strada da percorrere rimane quella di assumere valida l'ipotesi di omogeneità spaziale dei parametri, consapevoli che questa assunzione induce un'approssimazione nei risultati che è necessario quantificare.

Si sono dunque verificate quali sarebbero le conseguenze di una tale assunzione sempre nella stessa area di interesse; il caso di riferimento è rappresentato dalla calibrazione locale e validazione dei parametri del modello su ogni singolo bacino. Si è voluto poi procedere alla contemporanea calibrazione e validazione del modello su due e poi su tutti e tre i bacini considerati mantenendo esattamente gli stessi parametri. I risultati della calibrazione mostrano un netto peggioramento della bontà del modello nel passaggio alla considerazione di più bacini contemporaneamente. Il comportamento del modello in validazione non è ben definito; in caso di "data scarcity" ci si attenderebbe che, all'aumentare del numero di bacini coinvolti nell'analisi, aumentino anche le informazioni utili per ricondursi a un modello più robusto. Disponendo almeno di cinque anni di dati giornalieri non si dovrebbe ricadere in questo caso e ci si attenderebbe una netta diminuzione anche del valore dell'efficienza calcolata in fase di validazione. Nella realtà si nota che il valore medio dell'efficienza calcolata in validazione non cala all'aumentare del numero di bacini considerati.

Risultati rassicuranti riguardano il caso di bacini considerati non strumentati; si nota infatti come, per ottenere una prima simulazione idrologica di buona qualità si possa applicare un modello afflussi-deflussi i cui parametri siano quelli derivanti dalla calibrazione del modello in un bacino strumentato posto nelle vicinanze di quello di interesse. Sicché i risultati ottenuti per il caso di bacini non strumentati sono incoraggianti, nella seconda parte di questo lavoro di tesi ci si concentra su un'analisi di tipo regionale fondata sull'ipotesi di invarianza spaziale dei parametri, in riferimento alla quale si è condotta una verifica di dettaglio del livello di approssimazione introdotto.

PARTE II:  
ANALISI REGIONALE DELL'IPOTESI DI OMOGENEITÀ  
PARAMETRICA

---



# Introduzione

Nel corso degli anni sono stati sviluppati numerosi modelli idrologici, strumenti ampiamente adottanti non solo in ambito scientifico ma anche per applicazioni pratiche; nella realtà però la maggior parte delle applicazioni ingegneristiche si scontra con un problema comune: la scarsità o mancanza totale di osservazioni di portata. Il ricorso alla simulazione idrologica nel caso di bacini scarsamente strumentati o non strumentati ha stimolato l'interesse della comunità scientifica che ha posto questo tema al centro dell'attenzione (*Sivapalan, 2003*); la mancanza di misure di portata rende infatti impossibile la calibrazione locale dei parametri del modello considerato ed è dunque necessario ricorrere ad alternative. L'approccio solitamente scelto soprattutto nel caso di modelli afflussi-deflussi di tipo concettuale è il trasferimento dei parametri, che vengono calibrati sui bacini strumentati appartenenti ad una regione che può essere considerata idrologicamente simile, al bacino non strumentato, adottando differenti tecniche; questo processo viene definito come "regionalizzazione dei parametri". Esistono numerosi metodi per effettuare la regionalizzazione dei parametri (*Bloschl, 2005*), ma solitamente si segue una procedura composta da due fasi (*Parajka et al. 2007a*): si calibra il modello nei confronti delle serie di portata registrate nei bacini strumentati e poi si trasferiscono i parametri, adottando relazioni matematiche particolari, al bacino non strumentato oggetto di interesse.

Il vero e proprio trasferimento dei parametri viene effettuato sfruttando equazioni di regressione tra ciascun parametro del modello ed attributi geo-morfo-climatici propri dei bacini (strumentati e non) appartenenti alla regione studiata (*Yadav, 2007; Wagener, 2006; Parajka et al., 2005*). Si può infatti assumere che i parametri di un modello siano strettamente correlati agli attributi fisici di ciascun bacino, dunque si sviluppano equazioni che mettono in relazione ciascun parametro del modello, considerato indipendentemente dagli altri, con le caratteristiche fisiche e misurabili dei bacini. *Parajka et al. (2005)* ad esempio, considerano attributi che descrivono la topografia dell'area (area e quota caratteristica dei bacini oltre alla pendenza media), l'organizzazione del reticolo idrografico (densità di drenaggio), la geologia, l'uso del suolo, le condizioni climatiche medie dell'area, la presenza di specchi d'acqua dolce che provocano un'attenuazione delle portate, oltre ad altri attributi. Alcuni parametri possono risultare strettamente legati a particolari caratteristiche fisiche, ma non di rado si verifica il contrario: la correlazione tra i parametri del modello scelto e le proprietà dei bacini è molto debole (*Yadav et al., 2007; Bloschl, 2005*). *Wagener e Wheeler (2006)* ricordano che questo approccio risente delle incertezze legate alla scelta del modello idrologico più adatto per descrivere i fenomeni che accadono su tutti i bacini, oltre all'incertezza strutturale del modello. *Bloschl (2005)* spiega invece la debolezza delle correlazioni sottolineando che spesso si ricorre all'utilizzo di proprietà misurabili dei bacini che, pur essendo molto importanti, possono influenzare non in maniera significativa la formazione di deflusso. I problemi possono sorgere anche dalla scelta errata della relazione matematica che esprime il legame tra le proprietà fisiche ed i parametri del modello. Inoltre l'incertezza associata ai valori dei parametri può essere così

elevata da confondere le relazioni esistenti tra i parametri e gli attributi dei bacini (*Gottschalk, 2002*).

In letteratura sono state proposte analisi di regionalizzazione dei parametri alternative. Si può scegliere ad esempio di trasferire al bacino non strumentato l'intero set di parametri che deriva dalla calibrazione del modello su bacini che si trovano nelle sue immediate vicinanze, ipotizzando in questo modo che la vicinanza spaziale sia la misura che esprime la similitudine tra i bacini. Vandewiele and Elias (*1995*) si rifanno infatti a questo principio andando a paragonare due tecniche di regionalizzazione differenti: la prima basata sul kriging mentre la seconda prevede il trasferimento dei parametri che sono stati calibrati sul bacino più prossimo a quello non strumentato. Sempre su questa scia sono stati sviluppati diversi tentativi.

Merz e Blöschl (*2004*) paragonano invece otto differenti tecniche di regionalizzazione sviluppando il loro studio su 308 bacini che si trovano sull'intero territorio austriaco; tra i vari metodi analizzati considerano anche la possibilità di adottare un set di parametri costituito dal valor medio dei parametri del modello calibrati su tutti i bacini considerati, oppure considerano anche un insieme di parametri ottenuto come media tra i parametri ottimizzati sui bacini identificati dalle sezioni fluviali che si trovano immediatamente a monte e a valle di quella non strumentata oltre alla classica procedura di regionalizzazione dei parametri che ricorre all'uso di regressioni multiple tra i parametri del modello e le caratteristiche fisiche dei bacini.

Parajka et al. (*2007a*) propendono invece per una calibrazione regionale nella quale i parametri del modello vengano calibrati simultaneamente su diversi bacini; sono infatti convinti che la calibrazione del modello idrologico sfruttando le informazioni che provengono da un certo numero di bacini porti all'individuazione di un insieme di parametri più robusto e dunque a simulazioni idrologiche, realizzate sui bacini non strumentati, affette da un minor grado di incertezza.

Lo studio proposto in questa sezione segue questa stessa idea e si vuole proporre come metodo alternativo al classico utilizzo di regressioni tra i singoli parametri del modello e le caratteristiche fisiche dei bacini.

# 1 Cross-calibrazione di modelli afflussi-deflussi a scala regionale

L'analisi descritta nella presente sezione punta l'attenzione verso i bacini non strumentati o scarsamente strumentati per verificare se, in mancanza di misure di portata, sia possibile parametrizzare in modo soddisfacente i modelli idrologici. Sono questi i casi nei quali non si può ricorrere alla classica calibrazione dei modelli afflussi-deflussi e solitamente si sceglie di optare per l'applicazione di lunghe procedure di regionalizzazione dei parametri. Il metodo usualmente scelto per affrontare queste problematiche consiste cioè nel trasferimento, tramite particolari tecniche, dei parametri che sono stati valutati su bacini strumentati e dotati di caratteristiche idrologiche simili a quelle del bacino non strumentato oggetto di interesse.

Nel corso della presente trattazione si vuole proporre un metodo alternativo alla classica regionalizzazione dei parametri. Per semplicità si sceglie di definire “target” ogni bacino scarsamente strumentato o non strumentato, mentre si indicano come “donatori” tutti i bacini per i quali si dispone di misure di portata e sui quali si può procedere alla calibrazione locale dei parametri.

La procedura che si va a descrivere è basata sull'adozione, per tutti i bacini donatori, di un unico set di parametri che non varia nello spazio. Lo schema proposto implica dunque che i parametri possano essere direttamente applicati al bacino non strumentato, senza necessitare di ulteriori trasformazioni. Per poter valutare l'applicabilità dell'idea, si va a sviluppare un approccio jack-knife che simula il caso della presenza di bacini non strumentati: si considera cioè, di volta in volta, che tra tutti i bacini presenti nell'area di studio e per i quali si dispone effettivamente di osservazioni idrometriche e meteorologiche ne esista uno solo sul quale non si hanno informazioni di portata. Questo bacino, per le analisi che seguono, viene dunque ipotizzato come non strumentato e rappresenta il “target” delle simulazioni idrologiche; tutti gli altri bacini, dotati di misure di portata, sono invece denominati “donatori”. Si simulano le dinamiche di bilancio idrologico del bacino “target” sfruttando parametri stimati unicamente attraverso informazioni derivanti dai bacini “donatori”.

Nello specifico sia  $n$  il numero di bacini presenti nell'area oggetto di studio e si identifichi il primo bacino “target”; la procedura di cross-calibrazione necessiterà quindi di  $n - 1$  successivi passaggi. Si ordinino i bacini “donatori” in base alla particolare misura che esprime la similitudine idrologica di questi nei confronti del bacino “target”. Si suddivida poi ciascuna serie temporale di portata in due distinte porzioni: una deputata alla calibrazione del modello ed una invece per la sua validazione.

Il primo passo della procedura prevede la calibrazione del modello afflussi-deflussi scelto sul bacino “donatore” più simile al “target”. Il secondo passo prevede poi di calibrare il modello idrologico sul primo e sul secondo “donatore” scelti in base all'ordine di similitudine stilato in

precedenza. La calibrazione realizzata al secondo passo considera dunque contemporaneamente i due bacini “donatori” ed adotta per loro il medesimo set di parametri. Si procede cioè all’ottimizzazione del modello in modo che le serie di portata simulate in corrispondenza delle due sezioni di chiusura dei bacini “donatori” considerati si discostino il meno possibile dalle serie di portata ivi osservate. In questo modo si applica l’ipotesi di omogeneità dei parametri nello spazio trattata anche nella sezione precedente del presente lavoro di tesi. Prevedendo la calibrazione simultanea del modello su due bacini donatori si sta esprimendo la convinzione che inglobando un numero maggiore di informazioni, anche derivanti da aree differenti, sia possibile pervenire alla calibrazione di un insieme di parametri più robusto. Con il terzo passo della procedura di cross-calibrazione regionale proposta, la calibrazione del modello avviene simultaneamente sui primi tre bacini “donatori” scelti dalla medesima lista di similitudine e si procede in questo modo fino alla contemporanea adozione di tutti i bacini “donatori” in fase di calibrazione.

Il set di parametri identificato ad ogni passo viene poi applicato al bacino “target” senza ricorrere ad alcuna tecnica di trasferimento; si verifica la bontà del modello così ideato valutandone la capacità predittiva sia sul periodo scelto per la calibrazione che sul periodo deputato alla validazione. L’intero processo di cross-calibrazione viene ripetuto su ciascun bacino presente nell’area di studio, considerando in questo modo un bacino “target” alla volta, proprio come stabilito dalla tecnica jack-knife.

La tecnica di cross-calibrazione così ideata vuole verificare quale sia il giusto compromesso tra lo svantaggio derivante dall’assunzione di parametri omogenei nello spazio e il beneficio di poter aggiungere nuove informazioni ad ogni nuovo passo della cross-calibrazione; si vuole comprendere dunque fino a che punto, incrementando di volta in volta il numero di bacini considerati contemporaneamente, si giunga ad una stima maggiormente robusta dei parametri del modello e dunque se sia possibile produrre simulazioni idrologiche più attendibili.

L’analisi presentata può essere applicata in linea di principio a qualunque area di studio, adottando un generico modello afflussi-deflussi. La scelta poi dei bacini, le cui informazioni vengono considerate per la calibrazione dei parametri che sono poi trasferiti sul bacino non strumentato, è basata su una misura di similitudine; si tende infatti a considerare i bacini come donatori in ordine di similitudine decrescente con il sito di interesse. Particolarmente delicata è proprio la scelta della misura di similitudine tra i bacini, in base alla quale si stila l’ordine con il quale si scelgono i bacini “donatori” ad ogni passo della procedura di cross-calibrazione.

Le scelte che vengono effettuate in questa sede sono motivate dalla necessità di adottare un modello idrologico caratterizzato da buone prestazioni, di avvalersi di un’area di studio sufficientemente ampia e sulla quale siano disponibili dati di buona qualità. La qualità dei dati influenza infatti la bontà delle simulazioni in modo sostanziale, ancor più della scelta di un modello adeguato (*Perrin et al. 2001*).

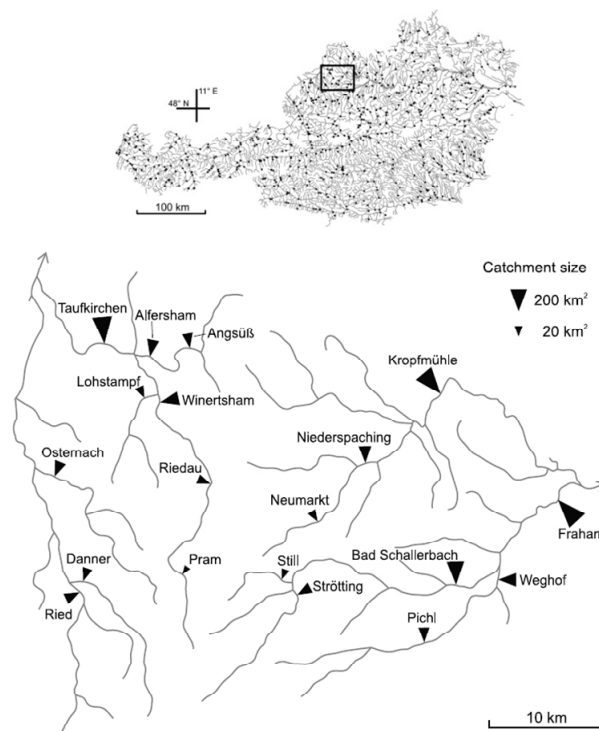


## 2 Modelli idrologici e aree oggetto di studio

I risultati riportati nella presente trattazione sono il frutto di analisi effettuate riferendosi a due regioni completamente differenti per morfologia, clima ma anche per qualità dei dati. Da un lato infatti si è voluta testare l'applicabilità della metodologia proposta su un'area vasta, nella quale sono riconoscibili un elevato numero di bacini dotati di lunghe serie di osservazioni idroclimatiche: la regione dell'Innviertel, localizzata nella parte nord dell'Austria. Dall'altro lato, poi, si sono cercate conferme applicando la medesima metodologia ad un'area caratterizzata da un limitato numero di osservazioni di portata, un'area che racchiude i bacini dei fiumi Enza, Secchia, Panaro e Reno collocata cioè nell'Italia centro-settentrionale.

Si presentano poi in questo capitolo i modelli concettuali considerati per realizzare le analisi regionali su entrambe le aree di studio e da ultimo si descrive quale misura viene adottata per stabilire la similitudine tra i bacini.

### 2.1 La regione dell' Innviertel



**Figura II- 1:** Regione dell'Innviertel, localizzata nella parte nord dell'Austria. (fonte: *Skoien and Bloeschl, 2007*)

I dati analizzati in questo studio derivano da un'estesa base dati dell'idrografia austriaca. Ci si focalizza in particolare sulla regione denominata Innviertel, nel nord ovest dell'Austria (

Figura II- 1) e già ampiamente studiata da parte di altri autori (*Skoien and Bloeschl, 2007*). Si è scelto di concentrare lo studio su quest'area, di dimensioni ragguardevoli (approssimativamente di 1500 km<sup>2</sup>), poiché è dotata di ben 19 stazioni di monitoraggio di portata che sottendono bacini le cui dimensioni variano dai 14 ai 375 km<sup>2</sup>. E' dunque una regione che offre la possibilità di testare l'idea alla base dello studio, disponendo di un numero ragguardevole di bacini, anche annidati l'uno all'interno dell'altro, di varie dimensioni. E' caratterizzata da condizioni di deflusso medio annuo complessivamente piuttosto uniformi, che varia infatti tra i 350 e i 510 mm/anno. La presenza di differenze marcate nella geologia, composta in larga parte da un'alternanza di morene, argille, marne e depositi ghiaiosi che riempiono le valli, porta ad avere un certo grado di variabilità anche nelle condizioni idrologiche che si riscontrano all'interno di ciascun bacino. Si tratta dunque di un'area sufficientemente ampia, caratterizzata da condizioni geo-morfologiche tali da poter ben rappresentare la base per uno studio di tipo regionale.

Per questa zona si dispone di osservazioni idrometriche e meteorologiche che si estendono su un periodo di oltre 30 anni, un periodo cioè decisamente lungo, sul quale è possibile realizzare simulazioni affidabili; precisamente si dispone di dati giornalieri di pioggia cumulata, di temperatura media e di portata sull'intervallo temporale che si estende dal 01/01/1976 al 31/12/2008 oltre che di stime di evapotraspirazione potenziale determinate mediante il metodo Blaney-Criddle modificato (*Parajka et al., 2005*) sempre sul medesimo intervallo temporale .

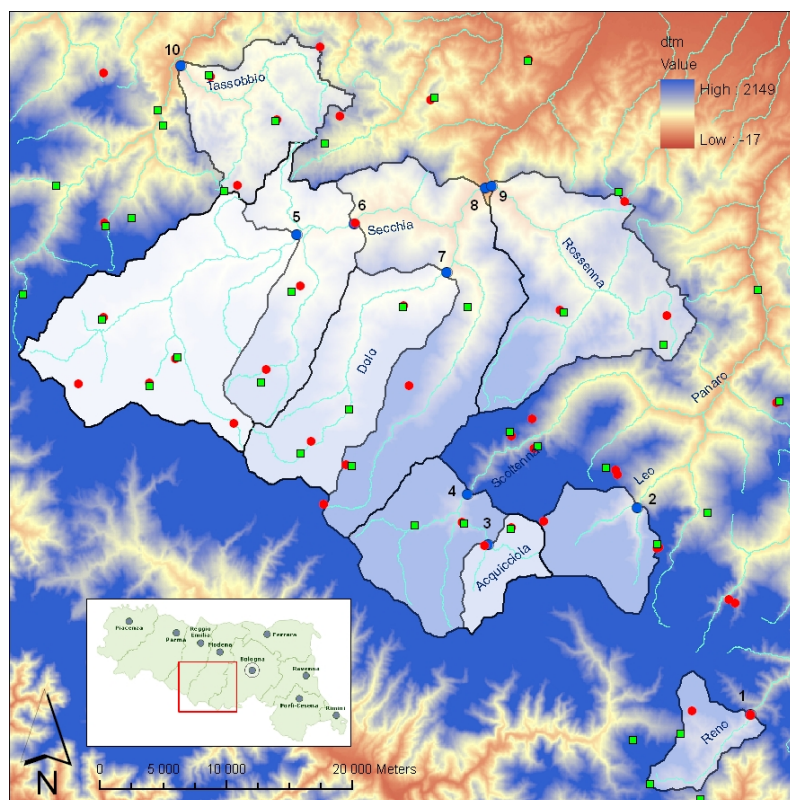
codice	Nome della stazione	area (km <sup>2</sup> )
204859	Taufkirchen	303.3
204834	Alfersham	81.3
204826	Angsüß	64.1
206268	Lohstampf	39.3
206326	Winertsham (Steg)	128.1
204768	Osternach	68.6
204784	Riedau	60
206805	Danner	55.6
206102	Pram	14.2
204735	Ried im Innkreis (Ort)	69.3
205054	Kropfmühle	312.2
206169	Niederspaching	101.9
205047	Neumarkt im Hausruckkreis	25.8
204990	Still	19.8
204974	Strötting	54
204958	Pichl bei Wels	66.8
204966	Weghof	116.9
205039	Fraham	375.2
205021	Bad Schallerbach	184

**Tabella II- 1:** Indicazione delle sezioni di chiusura considerate all'interno della regione di studio austriaca.

## **2.2 Il contesto italiano: i bacini dei fiumi Enza, Secchia, Panaro e Reno**

I fiumi Enza, Secchia, Panaro, fiumi appenninici ed affluenti di destra del fiume Po, ed il fiume Reno sono monitorati in diverse sezioni. Si sceglie di studiare unicamente i sottobacini di questi fiumi i cui deflussi risentano limitatamente della pressione antropica e le cui stazioni idrometriche riportino osservazioni di portata su un intervallo reputato sufficientemente lungo. Tali condizioni portano a limitare lo studio ai 10 bacini rappresentati in Figura II- 2. In particolare per il fiume Enza, si sceglie di analizzare solo il bacino del torrente Tassobbio e sul fiume Reno si considera unicamente la sezione idrometrica di Pracchia che chiude il bacino di questo fiume nella sua parte collinare. In corrispondenza di tali sezioni si dispone di osservazioni di portata giornaliera unicamente a partire dall'anno 2003, eccezion fatta per il fiume Reno sul quale le osservazioni hanno inizio nel 2002.

L'intera area così individuata è caratterizzata da condizioni morfo-climatiche non uniformi. Il bacino del torrente Rossenna, quello del torrente Tassobbio e quello del fiume Reno in particolare si distinguono rispetto agli altri; non comprendono infatti una porzione caratterizzata dalla presenza di alte montagne propria invece degli altri bacini. Il bacino del Rossenna e quello del Tassobbio poi sono caratterizzati da temperature medie molto più alte rispetto a quelle proprie del bacino del fiume Reno e da versanti più dolci. Il regime delle piogge poi è molto vario su tutta l'area di studio: una piovosità elevata è infatti presente nelle zone più prossime al crinale appenninico e man a mano che si scende verso la pianura il tasso di piovosità diminuisce. Tutti i corsi d'acqua appenninici alternano fortissime magre estive a piene primaverili e soprattutto autunnali da moderate a imponenti. Nel prosieguo ci si riferirà all'area interessata da questi bacini come alla "regione Emilia".



**Figura II- 2:** bacini oggetto di studio presenti nell'area compresa tra i fiumi Enza, Secchia, Panaro e Reno; in blu sono indicate le stazioni idrometriche, in rosso le stazioni termometriche e in verde quelle pluviometriche.

Codice	Corso d'acqua	Stazione	Area (km <sup>2</sup> )
1	Reno	Pracchia	40
2	Leo	Fanano	64.6
3	Acquicciola	Fiumalbo	18
4	Scoltenna	Pievepelago	130
5	Secchia	Gatta	233
6	Secchia	Ponte Cavola	348
7	Dolo	Ponte Dolo	134
8	Secchia	Lugo	693
9	Rossenna	Rossenna	187
10	Tassobbio	Compiano	101

**Tabella II- 2:** Indicazione delle sezioni di chiusura considerate all'interno della regione costituita dai bacini dei fiumi Enza, Secchia, Panaro e Reno.

Si dispone di misure di portata giornaliera sul periodo che intercorre tra il 01/01/2003 e il 31/12/2009 in corrispondenza delle sezioni indicate in Tabella II- 2; a completare il quadro delle osservazioni disponibili nella regione Emilia concorrono i dati di precipitazione cumulata e di temperatura media giornaliera osservate nelle stazioni meteorologiche indicate in Figura II- 2 che si riferiscono al medesimo intervallo temporale.

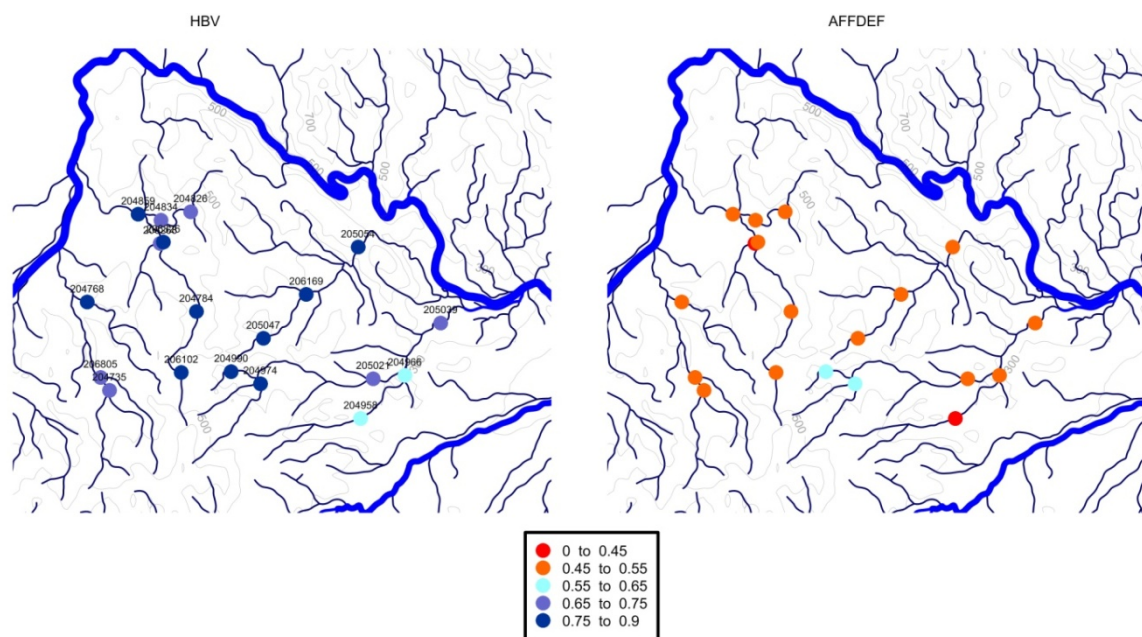
## 2.3 Scelta del modello afflussi-deflussi

La procedura di cross-calibrazione riportata al paragrafo 1.1 è un'analisi che necessita di un numero molto alto di simulazioni; la scelta del modello afflussi-deflussi da adottare non può quindi ricadere sulla tipologia costituita dai modelli spazialmente distribuiti, che

necessiterebbero di tempi di calcolo onerosi. Si è quindi optato per la classe dei modelli concettuali a parametri concentrati e si sono valutate anzitutto due differenti opzioni: una versione concentrata del modello AFFDEF presentato nel presente lavoro di tesi e la versione di HBV adottata da Parajka et al. (2007).

Il modello AFFDEF, spazialmente distribuito, può essere ridotto ad una versione concentrata semplicemente adottando, come cella di discretizzazione, una unica cella di dimensione pari all'area del bacino di interesse. Si sollecita dunque il sistema fornendogli in ingresso non più le serie di osservazioni puntuali di precipitazione e temperatura, ma quelle medie areali. Con una prima analisi si sono volute paragonare le prestazioni dei due modelli applicati ai bacini presenti nella regione dell'Innviertel. La valutazione della bontà di entrambi i modelli è stata effettuata in base alle prestazioni riportate in fase di calibrazione locale; in Figura II- 3 si mostrano i risultati ottenuti in termini di efficienza di Nash-Sutcliffe (*Nash and Sutcliffe, 1970*) e si nota come il modello HBV sia nettamente più adatto rispetto alla versione concentrata di AFFDEF, giustificandone quindi l'adozione nella trattazione regionale.

Poichè si è scelto di studiare il comportamento dei bacini emiliani, dotati di serie limitate di osservazioni di portata, si preferisce valutare la bontà della procedura di cross-calibrazione a scala regionale anche mediante un modello parsimonioso, HYMOD (*Boyle, 2000*), in modo da limitare l'incertezza associata ai parametri.



**Figura II- 3:** Prestazioni dei modello HBV e AFFDEF (nella versione concentrata) valutate in termini di efficienza di Nash-Sutcliffe.

## 2.4 Il modello HBV

Il modello HBV è stato originariamente sviluppato presso l'Istituto di Meteorologia ed Idrologia Svedese (SMHI) a partire dal 1971 e negli anni ha subito ulteriori modifiche perché è stato rielaborato sia degli stessi autori che lo hanno ideato, sia da altri studiosi che lo hanno adottato e adattato alle proprie necessità. Bergstrom (1991), consapevole dell'esistenza di numerosi modelli estremamente complicati, la cui complessità non è giustificata dalla qualità delle simulazioni che producono, propende per lo sviluppo di un modello concettuale che necessiti di pochi dati in ingresso e che richieda un numero ragionevolmente limitato di parametri. Esistono numerose versioni del modello HBV; una delle prime (Bergstrom, 1976) prevedeva unicamente la simulazione del deflusso superficiale ma negli anni sono state apportate numerose modifiche a ciascuna delle sue componenti; in seguito a studi approfonditi delle dinamiche proprie del sottosuolo si è proceduto, ad esempio, alla trasformazione dell'espressione della funzione di risposta del bacino (Bergstrom and Sandberg, 1983) così da meglio simulare gli effetti di falda. Come riportato in Bergstrom (1992), HBV è stato applicato, adottandone diverse versioni volta per volta, in bacini caratterizzati da svariati regimi idrologici e dislocati in oltre 40 diversi paesi del mondo; si sono realizzate simulazioni idrologiche con HBV in bacini che comprendono ghiacciai o sono caratterizzati dalla presenza di neve (ad esempio nelle Alpi, Himalaya, Nuova Zelanda, Groenlandia e Nord America), bacini artici (in Scandinavia e in Alaska), e bacini localizzati in ambienti umidi o semi-aridi che si trovano alle basse latitudini (come Panama e la Tunisia). Le dimensioni caratteristiche delle aree studiate poi sono estremamente variabili, si ricordi ad esempio l'applicazione di HBV (Bergström and Carlson, 1994; Graham, 1999) all'intero bacino che drena nel mar Baltico. La simulazione di aree così estese si può realizzare pensando che un bacino di vaste dimensioni è la somma di numerosi piccoli bacini; HBV infatti può anche essere adottato nella sua forma semi-distribuita, pensando di suddividere cioè il bacino oggetto di studio in sotto-bacini, oppure in fasce altimetriche o ancora in aree che si differenziano perché caratterizzate dalla presenza di laghi o di una particolare tipologia di vegetazione.

Solitamente si adotta un passo temporale giornaliero (riferendosi dunque a dati giornalieri sia di pioggia che di temperatura e a stime di evapotraspirazione potenziale giornaliera o mensile), ma esistono particolari versioni di HBV che simulano con un passo molto più basso (Boeschl et al, 2008). La versione messa a punto da Bloeschl et al. (2008) presenta alcune peculiarità: si tratta di un modello che aggiunge alle caratteristiche salienti di HBV numerosi ulteriori accorgimenti, simula su un intervallo temporale di 15 minuti e viene adottato per la previsione di piene improvvise nel bacino denominato Kamp, situato nella parte settentrionale dell'Austria e di dimensioni ragguardevoli (1550 km<sup>2</sup>). Per poter descrivere accuratamente ciò che accade in questo bacino gli autori hanno applicato una versione a tutti gli effetti spazialmente distribuita del modello; hanno infatti suddiviso il bacino in 13 sotto-bacini, ciascuno dei quali viene ulteriormente frazionato in unità fondamentali di area pari a 1 km<sup>2</sup> sulle quali si procede al calcolo del deflusso superficiale. Nel corso della presente trattazione si farà ancora riferimento a questa particolare versione di HBV, sottolineandone le particolarità.

Si presenta ora la versione originaria del modello HBV schematizzata in Figura II- 4, così come studiata al SMHI, che necessita, come dati in ingresso, di valori di precipitazione e, nelle

aree in cui è prevista la presenza di neve, anche delle osservazioni di temperatura, oltre che delle stime di evapotraspirazione là dove non siano invece disponibili misure realizzate sul campo attraverso evaporimetri. Come già accennato, solitamente si ricorre ad un passo di calcolo giornaliero ma si possono adottare intervalli di calcolo inferiori, se si dispone di osservazioni

#### **2.4.1.1 Il modulo neve**

Il modulo che viene implementato solitamente in HBV per rendere conto dei processi di accumulo e scioglimento nivale è un approccio gradi-giorno, che considera unicamente la temperatura come indicatore della presenza o meno di precipitazione in forma solida. Si assume infatti che la precipitazione sia neve quando la temperatura dell'aria è inferiore ad un particolare valore di soglia (TT) e, per considerare gli errori di misura inevitabilmente commessi sul campo della quantità di neve caduta, si utilizza un fattore di correzione SCF. Lo scioglimento dello strato di neve accumulato al suolo avviene quando la temperatura dell'aria supera il valore TT e segue l'espressione:

$$Melt = C_{melt} \cdot (T - TT) \quad (2.1)$$

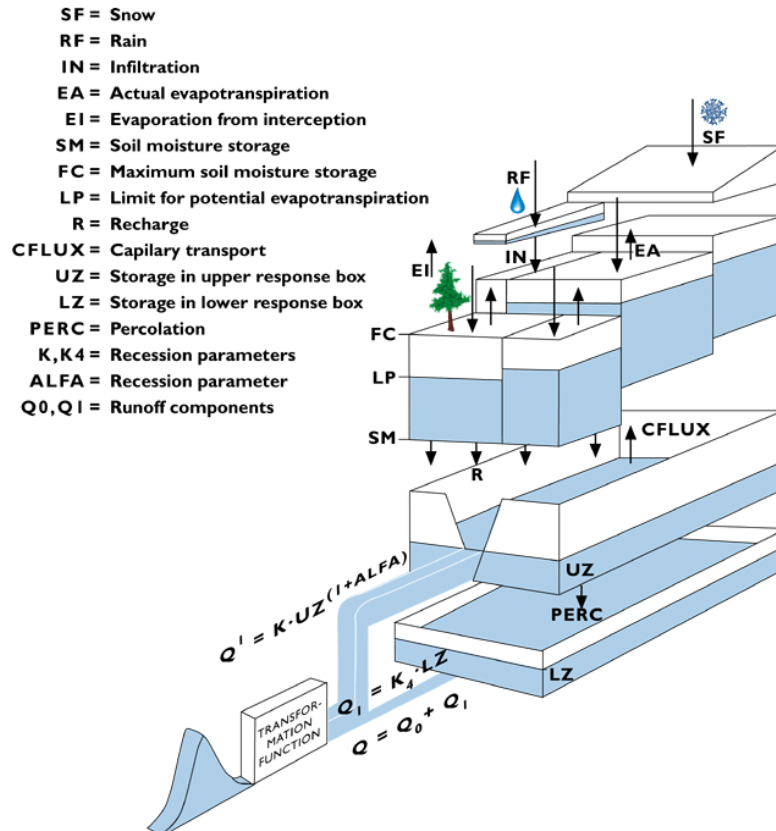
nella quale *Melt* rappresenta lo scioglimento nivale (mm/day),  $C_{melt}$  è il fattore gradi giorno (mm/°C/day) e TT è la temperatura di soglia (°C). Si assume poi che lo scioglimento effettivo della neve sia impedito fino a quando questo non supera una certa percentuale (ad esempio il 10%) del contenuto di neve nello strato al suolo. Questa è la più semplice formulazione possibile, che può essere in parte rielaborata e modificata con l'aggiunta di ulteriori parametri per considerare ad esempio la gradualità del passaggio tra neve e pioggia, oppure l'influenza della pioggia che, cadendo sullo strato nevoso, ne accelera lo scioglimento (come previsto in *Bloeschl et al, 2008*). Se poi si distinguono diverse zone all'interno del bacino (fasce altimetriche o zone caratterizzate da vegetazione differente), è necessario approntare un modulo neve differente per ciascuna di queste.

#### **2.4.1.2 Intercettazione**

Le prime versioni del modello HBV trascuravano il fenomeno dell'intercettazione di parte della precipitazione liquida o la consideravano come parte delle correzioni da effettuare alla neve attraverso il coefficiente SCF. In alcune versioni invece è previsto che solo nelle aree boscate si consideri l'effetto dell'intercettazione.

#### **2.4.1.3 Umidità del suolo**

Le dinamiche con le quali si provvede alla suddivisione dei deflussi superficiali da quelli profondi sono particolarmente complessi e necessiterebbero di modelli estremamente complicati, ma se l'obiettivo che ci si pone è quello di simulare la generazione di deflusso alla scala di bacino, allora si può semplificare il problema. Il ricorso ad un modello a serbatoio lineare porterebbe a simulazioni estremamente semplificate, dunque l'approccio previsto dall'HBV si basa sulla sua trasformazione grazie all'introduzione di un ulteriore concetto: si assume infatti che le varie zone del bacino siano dotate di capacità di immagazzinamento differenti, rappresentabili attraverso una distribuzione statistica.

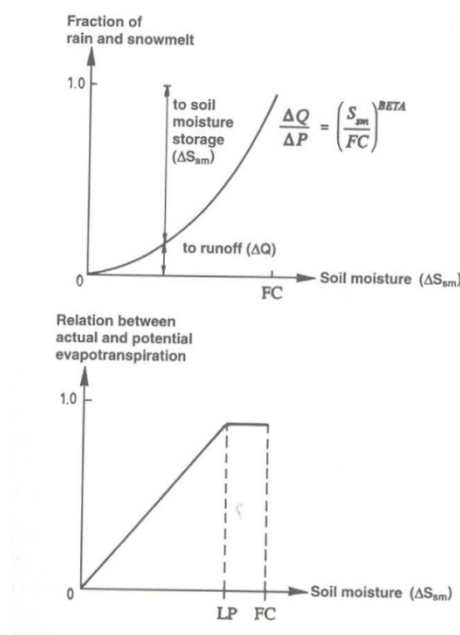


**Figura II- 4:** Struttura generale del modello HBV nella versione dello SMHI e pensato come applicato ad un sotto-bacino. (fonte: [http://www.macauley.ac.uk/hydalp/private/demonstrator\\_v2.0/models/hbv.html](http://www.macauley.ac.uk/hydalp/private/demonstrator_v2.0/models/hbv.html))

Ci si aspetta dunque che la pioggia (o lo scioglimento della neve) generi modesti contributi al deflusso quando il contenuto di acqua presente nel suolo è modesto, mentre sia responsabile di deflussi importanti quando il suolo ha un alto contenuto di umidità. Riferendosi all'equazione riportata in Figura II- 5, si indica con  $\Delta Q$  il contributo al deflusso superficiale, con  $\Delta P$  la precipitazione che raggiunge il suolo, con  $S_{sm}$  il contenuto di umidità presente nel suolo, con  $FC$  la capacità del serbatoio nel quale si accumula l'umidità e con  $BETA$  un ulteriore parametro che deve essere calibrato.

L'evapotraspirazione effettiva è poi calcolata come funzione dell'effettivo contenuto di umidità del suolo e dell'evapotraspirazione potenziale. Nel momento in cui il contenuto di umidità del suolo effettivamente presente all'interno del serbatoio supera il valore di soglia  $LP$ , l'evapotraspirazione effettiva uguaglia quella potenziale, altrimenti il valore effettivo risulta inferiore.





**Figura II- 5:** Schematizzazione del modulo responsabile della stima dell'umidità del suolo. FC, LP e BETA sono parametri empirici. (fonte: *Bergstrom, 1995*)

L'entità di acqua che esce dal sistema sottoforma di evapotraspirazione effettiva viene anzitutto sottratta dal serbatoio di intercettazione, quando questo esiste, in quantità pari a quella prevista dall'evapotraspirazione potenziale fino a quando il serbatoio stesso ne contiene, anche in forma di neve. Come già accennato, non sempre si prevede un modulo dedicato all'evapotraspirazione; talvolta semplicemente si decide di compensare la sua assenza agendo sul valore dell'evapotraspirazione potenziale: Johansson (1993) sceglie, ad esempio, di considerare l'evapotraspirazione potenziale delle zone boscate il 15% maggiore rispetto a quella delle altre zone.

#### 2.4.1.4 Funzione di trasferimento

Il modello HBV, messo a punto dal SMHI, prevede che l'eccesso di acqua proveniente dal serbatoio che regola la presenza di umidità nel suolo venga sommato alla pioggia che cade sugli specchi d'acqua attraverso una funzione di risposta di tipo concettuale. Il bacino cioè si comporta come se fosse dotato di due ulteriori serbatoi (i due invasi rappresentati in Figura II-4). L'invaso posto a quota inferiore si comporta a tutti gli effetti come un serbatoio lineare che rappresenta il comportamento della falda (è la risposta lenta del bacino) e si riempie tramite la percolazione di acqua proveniente dal serbatoio sovrastante ad una velocità costante *PERC*.

Se poi l'eccesso di acqua proveniente dal suolo è superiore alla percolazione, una parte di acqua viene immagazzinata all'interno del serbatoio superiore, il quale simula la risposta del bacino ad eventi di pioggia intensi (è la risposta veloce del bacino, responsabile della formazione delle piene). Nuovamente, le varie versioni di HBV presentano leggere differenze anche in questa parte di codice.

La versione alla quale si riferisce la Figura II- 4 (*Lindstrom et al., 1997*) prevede che l'apporto al deflusso proveniente dal serbatoio veloce sia rappresentata come  $Q_0 = K \cdot$

$UZ^{(1+ALFA)}$ , nella quale  $UZ$  rappresenta l'effettivo quantitativo di acqua immagazzinato, mentre  $K$  e  $ALFA$  sono due coefficienti; l'apporto di falda è invece dato da  $Q_1 = K_4 \cdot LZ$  nella quale  $K_4$  è l'ulteriore coefficiente, mentre  $LZ$  il contenuto del serbatoio. L'approccio classico invece adottato da Harlin (1992) e Bergstrom (1995), tra gli altri, prevede per il serbatoio superficiale sia la possibilità di sfioro, se il contenuto di acqua al suo interno supera la capacità del serbatoio stesso pari a  $L_{UZ}$ , sia uno scarico al fondo, mentre il serbatoio di falda viene modellato come in Lindstrom et al. (1997).

Qualunque sia la rappresentazione scelta per la trasformazione dell'umidità del suolo, le componenti rappresentanti la risposta veloce e lenta del bacino vengono sommate e ulteriormente filtrate attraverso una funzione di trasformazione triangolare. Si rimanda a Nester et al. (2011) come esempio di una versione di HBV che adotta una descrizione maggiormente articolata della trasformazione della pioggia netta in deflusso superficiale.

## 2.5 La versione del modello HBV adottata nella presente trattazione

Si sceglie di utilizzare una versione di HBV già ampiamente testata, conosciuta in letteratura ed adottata ad esempio da Parajka et al. (2007), derivante da quella originariamente proposta da Bergstrom (1976) e Lindstrom et al. (1997). Parajka et al. (2007) propongono un'applicazione di tipo semi-distribuito prevedendo la suddivisione di ciascun bacino in differenti fasce altimetriche all'interno delle quali si valuta sia il contenuto di umidità al suolo, sia l'accumulo e lo scioglimento di neve; nella trattazione che seguirà invece, si sceglie di adottarne una versione concentrata, trascurando quindi il frazionamento del bacino in differenti sotto-unità.

Si adotta un passo di calcolo giornaliero e vengono forniti in ingresso al modello, per ciascun bacino considerato, osservazioni di precipitazione cumulata giornaliera, temperatura ed evapotraspirazione potenziale media giornaliera.

Accumulo e scioglimento nivale sono rappresentati attraverso un modello gradi-giorno che prevede la distinzione della precipitazione  $P$  in pioggia,  $P_R$ , o neve,  $P_S$  secondo la:

$$\begin{aligned}
 P_R &= P \quad \text{se } T_A \geq T_R \\
 P_R &= P \frac{T_A - T_S}{T_R - T_S} \quad \text{se } T_S < T_A < T_R \\
 P_R &= 0 \quad \text{se } T_S > T_A \\
 P_S &= P - P_R
 \end{aligned} \tag{2.2}$$

nella quale  $T_A$ ,  $T_R$  e  $T_S$  sono rispettivamente la temperatura dell'aria media giornaliera e due temperature di soglia, superiore e inferiore che delimitano il campo delle temperature caratteristiche della precipitazione in forma liquida da quello tipico della precipitazione nevosa. Anche in questo caso si considerano i possibili errori di misura commessi nella rilevazione delle precipitazioni solide attraverso il coefficiente moltiplicativo  $SCF$ . Lo scioglimento nivale  $M$  è modellato esattamente allo stesso modo in cui avviene nel modello AFFDEF presentato nella

Parte I della presente trattazione; si prevede dunque l'utilizzo del fattore gradi-giorno  $DDF$  e del parametro denominato temperatura di fusione,  $T_M$ .

Il modulo "umidità nel suolo" rappresenta la generazione di deflusso superficiale e le variazioni di umidità nel suolo nel tempo attraverso la:

$$S_{SM,i} = S_{SM,i-1} + P_R + M - E_A \quad (2.3)$$

nella quale  $S_{SM,i}$  è il contenuto di umidità dello strato superficiale di suolo che controlla la generazione di deflusso superficiale all'istante  $i$ , mentre con  $E_A$  si indica l'evapotraspirazione effettiva. Il contributo che pioggia e neve apportano alla generazione di deflusso superficiale si denota con  $\Delta S_{UZ}$  e viene stimato secondo la classica schematizzazione:

$$\Delta S_{UZ} = \left( \frac{S_{SM}}{FC} \right)^{BETA} (P_R + M) \quad (2.4)$$

indicando con  $FC$  la capacità del serbatoio nel quale si accumula l'umidità, mentre  $BETA$  porta alla generazione di deflusso il carattere di non linearità. L'evapotraspirazione effettiva  $E_A$  risulta inferiore rispetto a quella potenziale  $E_P$  fino a quando lo strato superficiale di suolo si satura  $FC = S_{SM}$  e segue la classica trattazione:

$$\begin{aligned} E_A &= E_P \frac{S_{SM}}{LP} & \text{se } S_{SM} < LP \\ E_A &= E_P & \text{se } S_{SM} \geq LP \end{aligned} \quad (2.5)$$

essendo  $LP$  il parametro che esprime la capacità del suolo oltre la quale l'evapotraspirazione ammissibile eguaglia quella potenziale.

Con la trasformazione della pioggia netta in deflusso superficiale si vuole rappresentare la propagazione che avviene sulle pendici dei bacini ed è schematizzata ricorrendo a due serbatoi che rappresentano il comportamento di due differenti strati di suolo.  $S_{UZ}$  ed  $S_{LZ}$  siano rispettivamente le quantità di acqua immagazzinate nell'istante  $i$  all'interno del serbatoio che rappresenta la zona superiore e in quello profondo. La quantità  $\Delta S_{UZ}$  entra all'interno del serbatoio, che descrive ciò che accade nello strato di suolo superficiale, e segue tre diversi percorsi:

- percolazione, verso il serbatoio che rappresenta il comportamento della falda profonda, ad una velocità costante  $C_p$  e conseguente efflusso dal serbatoio stesso regolato dal parametro  $K_2$ ;
- scarico al fondo del serbatoio che rappresenta il comportamento dello strato di suolo più superficiale, simulato ricorrendo al parametro  $K_1$ ;
- si prevede per il serbatoio superficiale anche la possibilità di sfioro, modellato ricorrendo al parametro  $K_0$ , se il contenuto di acqua al suo interno supera la capacità del serbatoio stesso pari a  $L_{UZ}$ .

La somma dei deflussi così generati da entrambi i serbatoi viene indicata con  $Q_G$ ; questa viene ulteriormente trasformata per mezzo di una funzione di trasferimento triangolare che rappresenta la propagazione dei deflussi all'interno della rete drenante:

$$B_Q = B_{MAX} - C_R \cdot Q_G \quad \text{se } (B_{MAX} - C_R \cdot Q_G) \geq 1$$

$$B_Q = 1 \quad \text{altrimenti}$$
(2.6)

essendo  $B_{MAX}$  e  $C_R$  gli ulteriori parametri del modello.

Si sceglie di adottare come algoritmo di ottimizzazione per la presente trattazione l'algoritmo denominato SCE-UA, "shuffled complex evolution method" (Duan *et al.*, 1992); si tratta cioè di un algoritmo che rigetta un approccio basato sull'equifinalità e si prefigge invece la ricerca dell'insieme di parametri che rendano il modello scelto quello "ottimale". Il problema dell'esistenza di numerose regioni di attrazione, così come quello dell'esistenza di un alto numero di minimi locali (problema del quale si è discusso in precedenza nella parte I del presente lavoro di tesi) viene aggirato prevedendo due distinte fasi. Nella fase iniziale, a partire da una popolazione di punti campionati casualmente all'interno dello spazio dei parametri, vengono fatti evolvere indipendentemente un certo numero di complessi alla ricerca della soluzione ottimale; periodicamente poi si procede al completo rimescolamento della popolazione ed alla assegnazione dei punti ai diversi complessi. Così facendo i progressi effettuati da ciascun complesso possono essere condivisi con gli altri, consentendo una più rapida e robusta ricerca della soluzione, a patto che la popolazione di punti iniziale sia sufficientemente numerosa.

Analisi di sensitività realizzate dal gruppo di ricerca guidato dal prof. G. Bloeschl su bacini austriaci che appartengono a questa zona, hanno rivelato come sia possibile escludere dal processo di calibrazione i parametri  $T_R$  e  $T_S$ , per i quali le stesse analisi suggeriscono valori pari 2°C e -2°C rispettivamente. Rimangono quindi 13 parametri da calibrare.

Parametro	Descrizione	Dimensioni
<b>SCF</b>	Fattore correttivo	[-]
<b>DDF</b>	Fattore gradi-giorno	[L/T/t]
<b><math>T_R</math></b>	Temperatura di soglia (pioggia)	[T]
<b><math>T_S</math></b>	Temperatura di soglia (neve)	[T]
<b><math>T_M</math></b>	Temperatura di soglia (scioglimento)	[T]
<b>LP</b>	Soglia evapotraspirazione	[-]
<b>FC</b>	Altezza massima di umidità al suolo	[L]
<b>BETA</b>	Costante coinvolta nel calcolo di eccesso di umidità al suolo	[-]
<b>K<sub>0</sub>, K<sub>1</sub>, K<sub>2</sub></b>	Costanti di rilascio dei serbatoi che rappresentano il comportamento degli strati di suolo	[t]
<b><math>L_{UZ}</math></b>	Soglia che definisce la quota di sfioro	[L]
<b>Cp</b>	Velocità di percolazione	[L/t]
<b><math>B_{MAX}</math></b>	Base della funzione di trasferimento triangolare	[t]
<b><math>C_R</math></b>	Parametro di scala (funzione di trasferimento triangolare)	[-]

**Tabella II- 3:** Elenco dei parametri della versione del modello HBV che si adotta. ( $T_R$  e  $T_S$  non vengono calibrati)

## 2.6 Il modello HYMOD

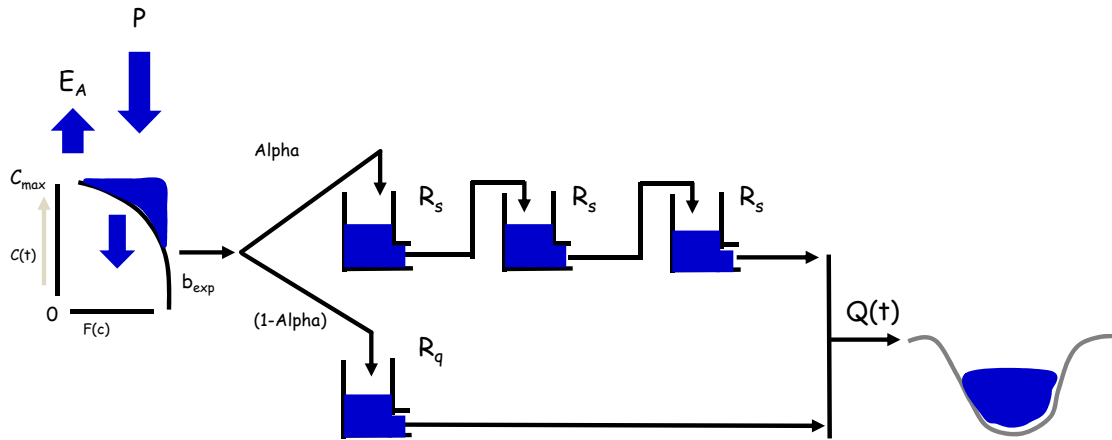


Figura II- 6: Schema di funzionamento del modello HYMOD

Il secondo modello adottato nell'analisi regionale che si propone in questo studio è il modello HYMOD, raffigurato in Figura II- 6, inizialmente introdotto da Boyle (2000); si tratta di un modello concettuale che per la sua semplicità, viene spesso adottato in numerose analisi (Vrugt et al., 2002; Vrugt et al., 2008; Wagener et al., 2001; Montanari and Toth, 2007 tra gli altri). E' infatti caratterizzato da 5 parametri concentrati, un numero decisamente limitato che lo rende estremamente parsimonioso essendo, al tempo stesso, in grado di ben adattarsi alle condizioni che si verificano in bacini completamente differenti.

Il concetto alla base del modello viene preso in prestito dagli studi di Moore (1985), il quale si era fatto persuaso della necessità di sviluppare modelli idrologici di complessità inferiore rispetto a quella propria dei modelli fisicamente basati, ma che rappresentassero in qualche modo la variabilità spaziale delle quantità idrologiche sul bacino. Per questo il suo approccio si basa sul principio della distribuzione di probabilità dei processi responsabili della formazione di deflusso. Moore sostiene che mentre i modelli fisicamente basati cercano di considerare l'effettiva distribuzione spaziale delle variabili idrologiche sul bacino, quelli che si basano sul principio di distribuzione probabilistica considerano unicamente la frequenza con la quale le variabili idrologiche assumono una certa grandezza. Per rendere però possibile l'applicazione di questo concetto è necessario che la rappresentazione della generazione di deflusso sia trattata con semplicità riferendosi, nel caso particolare, al concetto di "capacità di immagazzinamento". Si considera dunque che in ogni punto appartenente al bacino sia possibile imitare la variazione di umidità presente nel suolo adottando un serbatoio, di data capacità  $C$ , il cui contenuto viene incrementato dalla pioggia  $P$ , diminuito dall'evapotraspirazione  $E$ . Quando il contenuto di umidità supera la capacità del serbatoio si ha la generazione di deflusso. Un bacino dunque può essere considerato come la somma di numerosi serbatoi, ciascuno caratterizzato dalla propria capacità  $C$ , che interagiscono l'uno con l'altro. La variabilità nello spazio della capacità  $C$  di immagazzinamento di umidità è rappresentata dalla seguente funzione di distribuzione:

$$F(C) = 1 - \left(1 - \frac{C}{C_{max}}\right)^{b_{exp}} \quad \text{con } 0 \leq C \leq C_{max} \quad (2.7)$$

nella quale:  $F(C)$  è la probabilità cumulata della capacità di immagazzinamento di acqua,  $C_{max}$  [mm] è la massima capacità di immagazzinamento di acqua nel terreno. Il parametro  $b_{exp}$  [-] controlla il grado di variabilità spaziale della capacità di immagazzinamento sul bacino; nel caso particolare in cui  $b_{exp} = 0$ , il bacino sarebbe caratterizzato da capacità di immagazzinamento costante, mentre  $b_{exp} = 1$  implica che tale capacità varia uniformemente tra 0 e  $C_{max}$ .

La precipitazione che cade sul bacino può essere immagazzinata solo fino ad un certo limite, definito attraverso la funzione  $F(C)$ , oltrepassato il quale il terreno risulta completamente saturo, non può ricevere ulteriori apporti che vengono quindi destinati alla generazione di deflusso superficiale. La pioggia netta viene propagata mediante due serie di serbatoi lineari (tre, esattamente identici, che rappresentano la risposta veloce del bacino ed uno che raffigura la risposta lenta) posti in parallelo.

L'evapotraspirazione effettiva viene poi calcolata in accordo alla formula (2.8) e sottratta dal volume invasato all'interno del bacino  $W(t)$ .

$$E_A(t) = \left( 1 - \frac{\frac{C_{max}}{b_{exp} + 1} - W(t)}{\frac{C_{max}}{b_{exp} + 1}} \right) E_P(t) \quad (2.8)$$

Il modello dunque prevede la calibrazione di 5 parametri:

- $C_{max}$  [mm], la massima capacità di immagazzinamento di acqua nel terreno;
- $b_{exp}$  [-] che controlla il grado di variabilità spaziale della capacità di immagazzinamento sul bacino;
- $Alpha$  [-] il fattore che provvede alla separazione del deflusso tra le due serie di serbatoi;
- $R_s$  [t] che regola lo scarico al fondo dei serbatoi che riproducono la risposta veloce;
- $R_q$  [t] che regola lo scarico al fondo del serbatoio che riproduce la risposta lenta del bacino.

## 2.7 Misura di similitudine idrologica: distanza di Ghosh

La comunità scientifica si è a lungo interrogata sulla possibilità che bacini idrologicamente simili abbiano la tendenza a presentarsi l'uno in prossimità dell'altro. In particolare uno studio di Merz and Blöschl (2004) esamina la bontà di diversi metodi di regionalizzazione dei parametri ricorrendo all'applicazione di un modello concettuale su 308 bacini austriaci. Concludono che i metodi basati sulla vicinanza spaziale sono, nel loro caso, maggiormente efficienti rispetto a quelli basati sulla ricerca di attributi che esprimono la similitudine attraverso caratteristiche geo-morfo-climatiche. Una conferma di quanto provato in Austria si ha anche nel caso proposto da Oudin et al. (2008) i quali studiano 913 bacini situati in Francia ed affrontano il problema della regionalizzazione dei parametri adottando una regressione, ricorrendo alla regionalizzazione basata sulla vicinanza spaziale e realizzando un processo di regionalizzazione

basato sulla similitudine fisica. Dimostrano che la regionalizzazione basata sulla vicinanza geografica risulta essere l'opzione che porta ai risultati migliori nel loro caso.

Una possibilità per esprimere la similitudine idrologica tra diversi bacini è offerta dunque dal ricorso a misure di distanza spaziale, ma esistono diverse possibilità per esprimere la distanza tra due bacini idrologici (*Gottschalk et al., 2011*). Si può distinguere infatti la distanza euclidea tra le sezioni di chiusura dei due bacini, dalla distanza, misurata lungo l'asta principale, dei due bacini tra le sezioni di chiusura, dalla distanza euclidea tra i centri di massa dei due bacini, da un ultimo indicatore di distanza. Quest'ultimo, definito distanza di Ghosh (*Ghosh, 1951*), è un indicatore che esprime il valore atteso di tutte le possibili distanze misurate tra due differenti bacini.

Solitamente si ricorre alla stima della distanza tra due bacini definendola come la distanza tra i loro centri di massa, ma questo comporta alcuni svantaggi soprattutto quando si ha a che fare con bacini che sono annidati l'uno all'interno dell'altro. Due differenti bacini possono infatti avere il medesimo centro di massa, ma possono essere l'uno un sottobacino dell'altro, dunque se si adottasse questa misura di similitudine si incorrerebbe nel rischio di non riuscire a distinguere un bacino dal suo sottobacino.

Se si determina la distanza tra due bacini ricorrendo alla misura proposta da Ghosh, invece, non incorre in questo problema; si considera infatti un elevato numero di punti distribuito in modo omogeneo all'interno di ciascuno dei due bacini e si valutano le distanze tra tutte le possibili coppie di punti (ciascun punto appartiene a un diverso bacino). Un approccio tale è stato suggerito da Gottschalk (2006) proprio per nel caso si debbano considerare bacini annidati l'uno all'interno dell'altro.

Si sceglie dunque di adottare in prima battuta questa misura come indicatore della similitudine tra due bacini differenti, considerando dunque due bacini tanto più simili quando meno distanti l'uno dall'altro. Si ricorre all'uso del pacchetto denominato "rtop" (*Skoien et al., 2009*) che lavora in ambiente R per la stima della distanza di Ghosh tra tutti i bacini considerati nell'area di studio.

### **3 Problematiche legate alla scelta del periodo di calibrazione**

Il fulcro dell'intero studio è la ricerca di un giusto compromesso tra i benefici derivanti dalla possibilità di calibrare un modello afflussi-deflussi disponendo di un numero sempre crescente di informazioni e gli svantaggi dovuti all'utilizzo di un insieme di parametri invariati nello spazio su aree di dimensioni passo passo crescenti; la presente sezione esplora dunque quali siano le conseguenze dell'adozione dell'ipotesi di omogeneità spaziale dei parametri riferendosi al caso dell'analisi di tipo regionale.

Contestualmente si vuole verificare se la bontà della procedura di regionalizzazione proposta in questo studio sia intimamente legata al particolare periodo di calibrazione adottato; si replica dunque l'analisi regionale svariate volte scegliendo in ogni caso un periodo di calibrazione differente, non solo per lunghezza. Per simulare infatti la possibilità, che di sovente ricorre nella pratica, di disporre di bacini dotati di serie limitate di osservazioni per lo più riferite a intervalli temporali non concomitanti, si verificano anche gli effetti dell'utilizzo di periodi di calibrazione di pari lunghezza ma non contemporanei sui diversi bacini.

Per condurre un'analisi significativa per comprendere quale sia l'influenza del periodo di calibrazione sulla bontà dell'intera procedura di cross-calibrazione e cross-validazione a scala regionale, è necessario disporre di una lunga serie di dati; le stazioni idrometriche presenti nella regione dell'Innviertel riportano dati di portata su un periodo di lunghezza pari a 30 anni, mentre le osservazioni di portata registrate nei bacini emiliani si riferiscono al più a 7 anni, riportando, tra l'altro un cospicuo numero di dati mancanti. Inoltre, il numero di bacini presenti nella prima area di studio è sostanzialmente il doppio di quelli presenti nell'area Enza-Secchia-Panaro-Reno. Per questo si sceglie di concentrare le analisi che si riferiscono alla scelta del corretto periodo di calibrazione sull'area costituita dai 19 bacini situati nella regione denominata Innviertel rappresentata in

Figura II- 1. Per questi bacini si dispone di valori di precipitazione, temperatura ed evapotraspirazione giornaliera media areale per il periodo che si estende dal 01/01/1976 al 31/12/2008. Le stazioni idrometriche riportate in

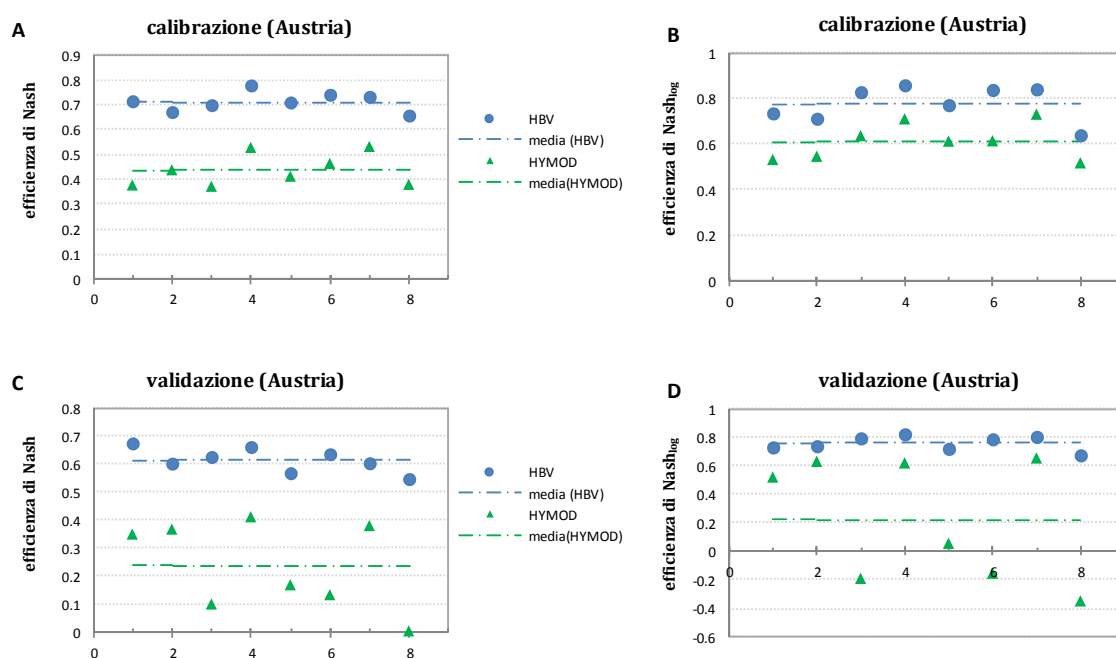
Figura II- 1 ed indicate in Tabella II- 1 sono dotate di misure giornaliere di portata sullo stesso intervallo temporale.

L'intero periodo sul quale sono disponibili dati idrologici e climatici viene suddiviso in tre porzioni di uguale ampiezza; la finestra centrale, che si estende dal 01/01/1987 al 31/12/1997, nella quale tutte le stazioni di misura riportano dati di portata, viene adibita alla calibrazione, mentre le due porzioni rimanenti vengono destinate alla validazione della procedura. Si parla dunque di cross-calibrazione riferendosi alla calibrazione regionale effettuata sempre



avvalendosi di dati che appartengono al periodo centrale (pur adottando di caso in caso periodi di lunghezza differenti), mentre con il termine cross-validazione si indica la procedura di verifica del modello caratterizzato da un particolare insieme di parametri, sul periodo costituito dai due intervalli temporali esterni.

Parajka et al. (2007) nello studio che conducono su 320 bacini presenti sull'intero territorio austriaco dimostrano come il modello HBV sia altamente prestazionale essendo in grado di ben simulare il comportamento idrologico della medesima area alla quale si rivolge l'interesse dello studio che si va ora a descrivere. Un paragone preliminare tra i modelli HBV e HYMOD calibrati (in senso classico) sull'intero periodo di calibrazione, e successivamente validati, sui soli 8 bacini che non sono annidati l'uno all'interno dell'altro conferma che il modello HBV è in grado di simulare più fedelmente i processi idrologici che avvengono nell'area identificata (si veda il confronto sia in fase di calibrazione che in fase di validazione tra i due modelli in Figura II- 7).



**Figura II- 7:** Indicatori di bontà dei modelli HBV e HYMOD riferiti alla fase di calibrazione (A e B) e validazione (C e D) effettuata nel senso classico sui soli 8 bacini austriaci non annidati.

Per questo motivo, anche se, in linea di principio, è possibile sviluppare la tecnica di regionalizzazione ricorrendo all'applicazione di qualsivoglia modello afflussi-deflussi, si sceglie di optare, per gli studi che si effettuano sulla regione dell'Innviertel, per l'applicazione a scala giornaliera del modello concettuale a parametri concentrati HBV.

Un periodo di riscaldamento pari a 305 giorni viene sempre adottato per consentire al modello afflussi-deflussi di svincolarsi dalla particolare scelta delle condizioni iniziali e si ottimizza il modello ricorrendo all'adozione di una funzione obiettivo  $L$  che focalizza la propria attenzione in ugual misura sia sulla corretta rappresentazione dei picchi dell'idrogramma che delle magre:

$$L = \left( 1 - \frac{\sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \overline{Q_{obs}})^2} + 1 - \frac{\sum_{i=1}^n (\ln(Q_{obs,i}) - \ln(Q_{sim,i}))^2}{\sum_{i=1}^n (\ln(Q_{obs,i}) - \ln(\overline{Q_{obs}}))^2} \right) / 2 \quad (3.1)$$

Il modello afflussi-deflussi scelto viene inoltre calibrato e validato in modo classico su ciascun bacino, così da costituire un punto di riferimento al quale paragonare i risultati dell'analisi regionale per giudicarne l'affidabilità.

Si può ottenere una buona riproduzione dei picchi di piena in riferimento al tempo, alla frequenza e al volume idrico coinvolto tramite l'uso di indici quadratici come l'indice di Nash-Sutcliffe (NS), mentre l'indice di Nash-Sutcliffe valutato invece sul logaritmo delle portate (NSlog) si concentra sulla rappresentazione del comportamento delle magre. Entrambi gli indici vengono adottati per valutare la bontà delle simulazioni effettuate. Si indica con  $Q_{obs,i}$  e con  $Q_{sim,i}$  il valore di portata osservata e simulata nell'intervallo temporale  $i$ , mentre  $\overline{Q_{obs}}$  e  $\ln(\overline{Q_{obs}})$  simboleggiano rispettivamente la media delle portate osservate non trasformate e log-trasformate.

Denominazione		Espressione matematica
NS	Efficienza di Nash-Sutcliffe	$NS = 1 - \frac{\sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \overline{Q_{obs}})^2}$
NSlog	Efficienza di Nash-Sutcliffe valutata sul logaritmo delle portate	$NSlog = 1 - \frac{\sum_{i=1}^n (\ln(Q_{obs,i}) - \ln(Q_{sim,i}))^2}{\sum_{i=1}^n (\ln(Q_{obs,i}) - \ln(\overline{Q_{obs}}))^2}$

Si è inoltre scelto di escludere dalle analisi due bacini, identificati dai codici 206805 e 204958, per i quali le simulazioni di calibrazione locale e validazione portano a risultati non soddisfacenti.

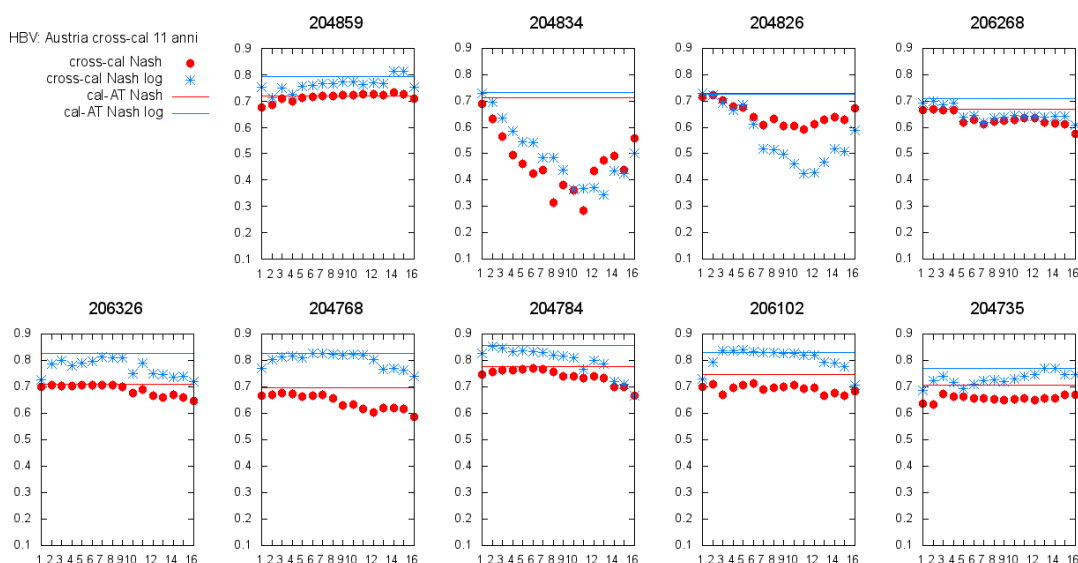
Nella trattazione che segue si presentano i risultati conseguiti nell'analisi regionale adottando periodi di calibrazione differenti.

### 3.1 Lunghezza del periodo di calibrazione

Per accertarsi che, mediante la calibrazione del modello idrologico, si identifichi un insieme di parametri che produce simulazioni che ben si adattano alle condizioni presenti sul bacino in qualunque periodo, Bergstrom (1991) suggerisce di adottare un periodo di calibrazione sufficientemente lungo. Disponendo di serie temporali eccezionalmente lunghe, si adotta, nel caso di seguito descritto, un intervallo pari a 11 anni (dal 01/01/1987 al 31/01/1997) sia per calibrare in modo classico il modello su ciascun bacino che viene considerato come target nelle analisi regionali, sia nelle cross-calibrazioni regionali. Il modello idrologico ottenuto si rivela in grado di ben rappresentare il comportamento di ciascun bacino sia nei periodi di morbida e magra che in quelli caratterizzati dalla presenza di piene; in Figura II- 8 (indicato con il tratto continuo), si nota infatti come la calibrazione effettuata su ciascun bacino in maniera classica porti a valori di efficienza di Nash-Sutcliffe valutati sulle portate osservate log-trasformate e su quelle non trasformate elevati. La procedura di cross-calibrazione a scala regionale prevede invece, per ciascun bacino "target", che la calibrazione del modello HBV avvenga sui bacini donatori e che solo successivamente il set di parametri identificato sia trasferito sul bacino

target. Si valuta dunque la bontà della procedura sia sul periodo di calibrazione che su quello di validazione.

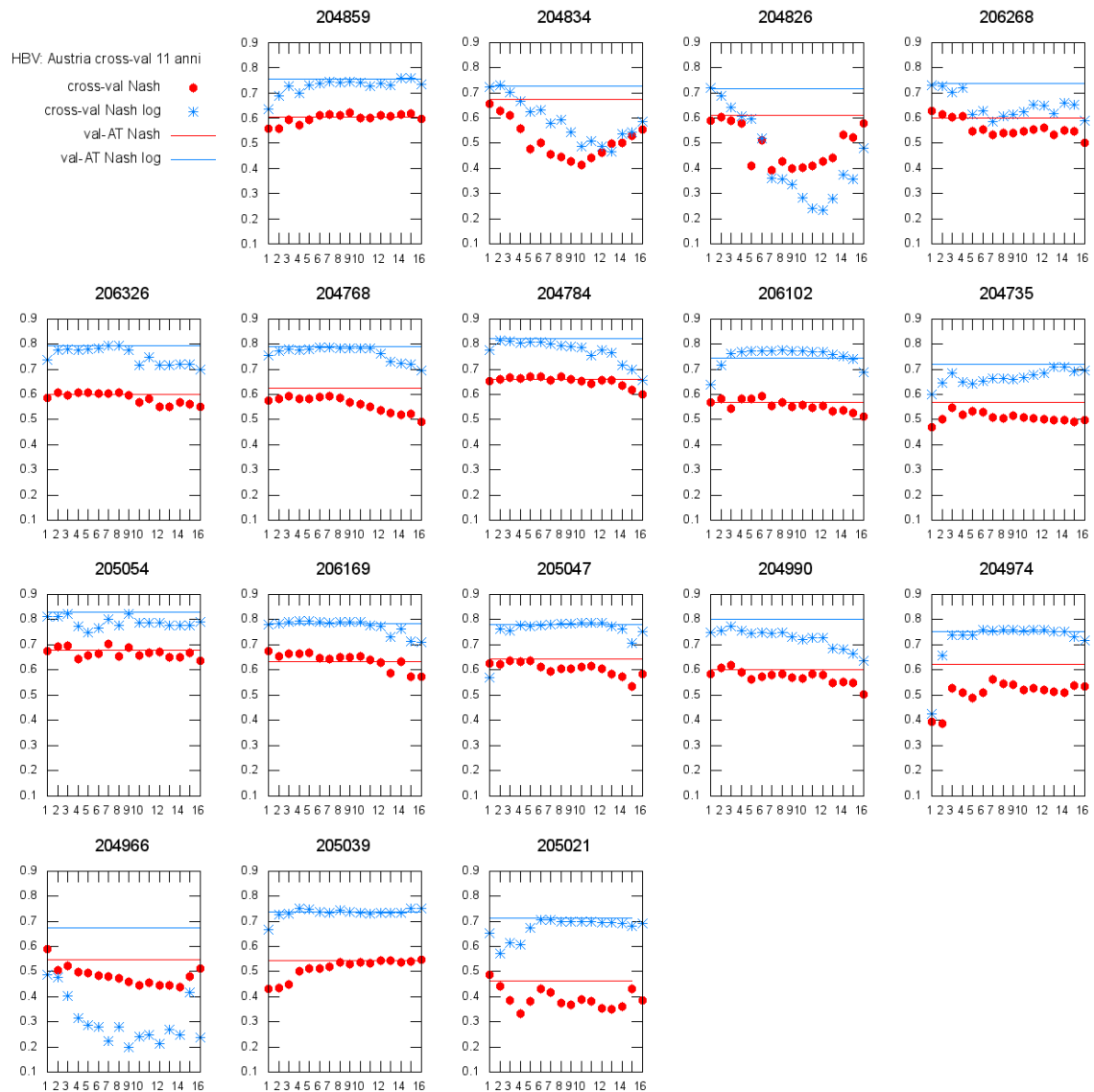
In Figura II- 8 sono riportati i risultati ottenuti nella procedura di cross-calibrazione; per una piena comprensione dei risultati ci si riferisce ora, a titolo di esempio, al caso del bacino identificato dal codice 204834. Il modello HBV calibrato localmente sul bacino 204834 si dimostra performante (il valore dell'efficienza di Nash-Sutcliffe calcolata sulle portate non trasformate è pari a 0.71 mentre quello valutato sulle portate log-trasformate è di poco superiore); questi risultati sono considerati come riferimento per verificare l'attendibilità della procedura regionale. I risultati della calibrazione regionale vengono rappresentati nello stesso grafico ricorrendo all'uso di pallini. La procedura di cross-calibrazione prevede che la calibrazione del modello avvenga di volta in volta su un numero crescente di donatori scelti secondo un particolare criterio di similitudine; avendo adottato come criterio di similitudine idrologica la vicinanza spaziale, il primo passo della procedura regionale prevede la calibrazione del modello sul bacino più prossimo al target. Il set di parametri così identificato viene applicato al bacino target, in questo caso il 204834, e si valutano le prestazioni del modello paragonando le portate simulate, con la calibrazione regionale, ed osservate nel bacino target sulla finestra temporale centrale, scelta per la calibrazione del modello, e poi anche sull'intervallo preposto alla cross-validazione. Nel caso del bacino 204834 il set di parametri derivante dalla calibrazione effettuata sul primo donatore porta a valori di efficienza valutati sul periodo di calibrazione poco inferiori rispetto a quelli che si ottengono attraverso una calibrazione classica.



**Figura II- 8:** Cross-calibrazione applicata alla regione dell'Innviertel adottando un periodo di calibrazione pari a 11 anni (01/01/1987-31/01/1997); risultati della calibrazione (locale, denominata cal-AT, e regionale, cross-cal) di 9 bacini espressi in termini di efficienza di Nash-Sutcliffe (Nash, in rosso) e di efficienza di Nash-Sutcliffe valutata sul logaritmo delle portate (Nash log, in blu).

Il secondo passo della procedura di cross-calibrazione consiste nell'utilizzare i due bacini più simili al 204834, nel calibrare dunque il modello, utilizzando un set di parametri costante su entrambi, facendo sì che riproducano al meglio contemporaneamente le condizioni idrologiche che si verificano sui due bacini donatori. Si trasferisce dunque l'insieme di parametri

identificato al bacino target e si valutano le prestazioni del modello sul periodo di calibrazione. Nel caso oggetto di attenzione si ottengono prestazioni inferiori rispetto a quelle del modello che è stato calibrato adottando come bacino donatore unicamente il più prossimo. La procedura di cross-calibrazione evolve poi al terzo passo considerando i tre bacini più simili al target, calibrando su di essi il modello ed applicando poi i parametri al bacino 204834. Le prestazioni del modello ottenuto adottando i parametri derivanti dalla contestuale calibrazione dei tre donatori subiscono nuovamente un calo. Di volta in volta dunque si considera un nuovo donatore, si incrementa così l'area interessata dall'ipotesi di omogeneità dei parametri e si valuta l'insieme di parametri che viene trasferito sul bacino "target" che si è ipotizzato come non strumentato.



**Figura II- 9:** Cross-validazione realizzata sugli anni 1976-1986 e 1998-2008 considerando di volta in volta un bacino target differente. Si valuta la bontà della procedura regionale paragonando i risultati ottenuti in cross-validazione (punti) e attraverso la validazione classica del modello (linea) su ciascun bacino target in termini di efficienza di Nash calcolata sulle portate osservate (in rosso) e sulle portate log-trasformate (in blu).

La procedura di cross-validazione consiste invece nella valutazione delle prestazioni del modello su un insieme di dati indipendente, mai adottato cioè in fase di calibrazione. Per questo ci si riferisce all'insieme di valori di portata che sono stati misurati sul periodo 1976-1986 e 1998-2008. I risultati della fase di cross-validazione sono riportati in Figura II- 9. Analizzando in dettaglio quanto accade al medesimo bacino del quale si è discusso in fase di cross-calibrazione, il 204834, si nota come, ipotizzandolo non strumentato e ricorrendo all'analisi regionale proposta, la qualità delle simulazioni effettuate sia decisamente inferiore rispetto al caso di bacino strumentato. Se si ipotizza cioè di non avere informazioni di portata sul bacino si può raggiungere una prima stima di buona qualità solo se si effettuano simulazioni idrologiche con un modello che viene sollecitato dall'insieme di parametri che è stato calibrato sul bacino a lui più prossimo. Man a mano poi che la cross-validazione evolve, considerando un insieme sempre più cospicuo di bacini donatori, si nota un netto peggioramento delle prestazioni del modello fino a raggiungere un minimo utilizzando 10 bacini donatori contemporaneamente, nel caso si valuti la bontà del modello puntando l'attenzione alla corretta rappresentazione dei picchi di portata, ed adottando 13 bacini donatori nel caso si sia interessati maggiormente alla rappresentazione delle magre. Le considerazioni effettuate fin ad ora per il caso della cross-calibrazione e cross-validazione avente come target il bacino 204834 si applicano a ciascuno degli altri 16 bacini che ricadono nell'area di studio.

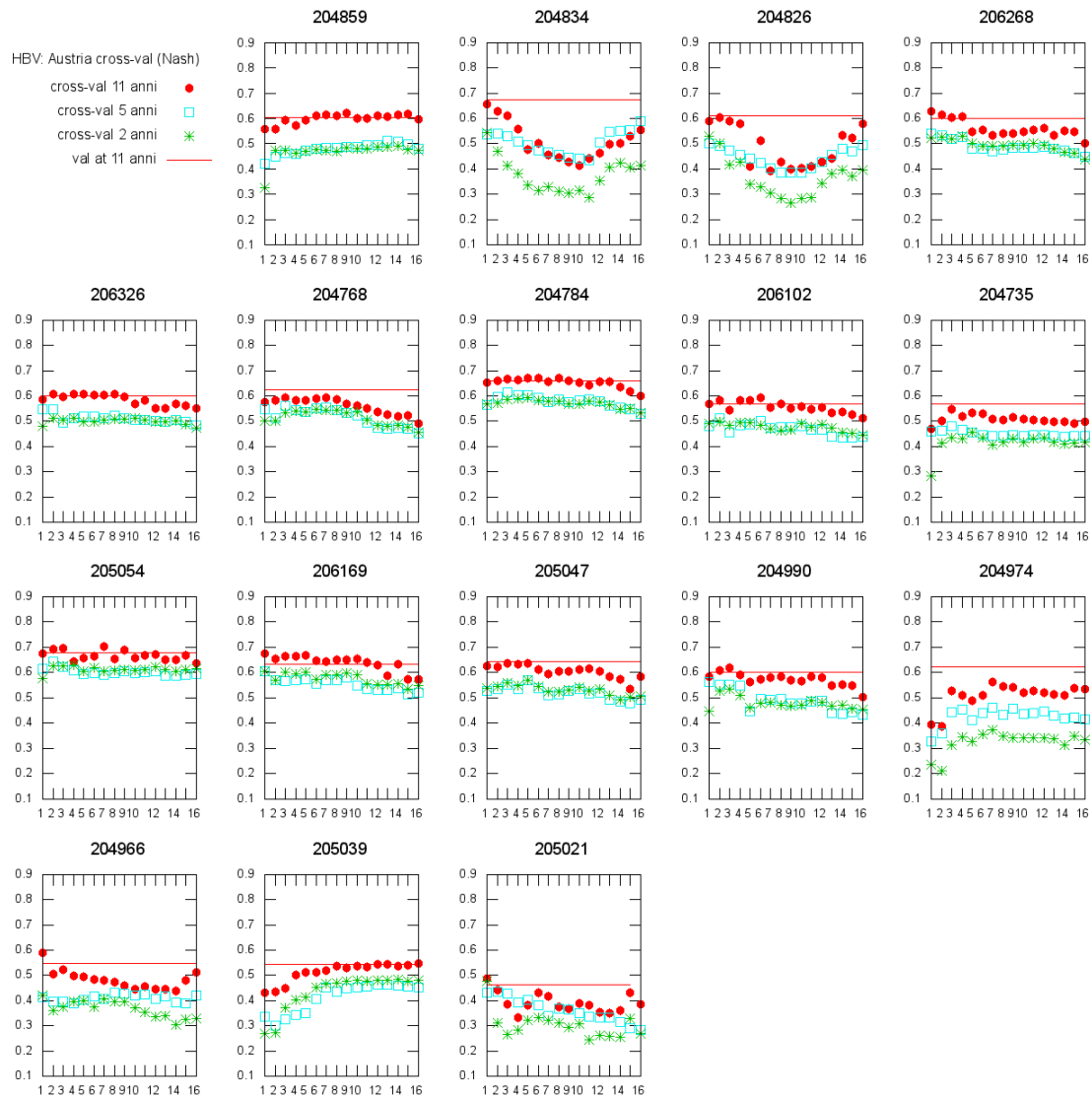
La procedura di calibrazione classica di un modello idrologico su un bacino dotato di misure di portata porta all'identificazione dell'insieme di parametri ottimale, dunque descrive nel migliore modo possibile le caratteristiche idrologiche del bacino di interesse. Mediante cross-calibrazione è possibile ottenere parametri che danno vita a modelli la cui bontà al più si accosta alle efficienze caratteristiche del modello calibrato classicamente.

Le potenzialità dell'analisi regionale si esprimono invece riferendosi al periodo di validazione (Figura II- 9). Si nota infatti che, nella maggioranza dei casi, attraverso l'analisi di tipo regionale è possibile realizzare modelli che, in quanto a bontà, si rivelano uguali o superiori a quelli determinati tramite la classica calibrazione. Il beneficio dato dall'incremento di informazioni in fase di calibrazione è risultato dunque superiore allo svantaggio derivante dall'utilizzo di parametri invariati su aree di dimensioni maggiori rispetto a quelle di un singolo bacino. Nel caso ad esempio del bacino codificato come 205039 si nota che l'analisi regionale porta al raggiungimento di efficienze di Nash-Sutcliffe (di entrambi i tipi) migliori rispetto al caso della validazione classica a partire dal momento in cui si considera una regione composta dagli 8 bacini ad esso più prossimi. In altri bacini si esprime un comportamento analogo (205047, 206169 e 206102 tra gli altri), indice del fatto che la calibrazione che viene effettuata adottando le informazioni provenienti da un alto numero di donatori porta a modelli di bontà paragonabile o anche più affidabili rispetto a quelli che derivano dalla calibrazione classica.

Sono però presenti anche casi in cui (si veda ad esempio i bacini 204834 e 204826) solamente i parametri derivanti dalla calibrazione del bacino ad essi più prossimo portano ad efficienze paragonabili a quelle del caso strumentato. Esiste poi anche il caso del bacino 204966 per il quale nessuna analisi regionale porta a modelli in grado di ben rappresentare contemporaneamente il comportamento del bacino nei confronti dei picchi di portate alte e delle morbide. Si vede dunque che nella maggioranza dei casi il ricorso alla procedura regionale che faccia uso di un numero alto di bacini donatori porta a risultati persino migliori rispetto a quelli

che si ottengono calibrando localmente e validando il modello nel caso strumentato; non è possibile però individuare un comportamento univoco e stabilire quale sia, ad esempio, il numero di bacini che porta al miglior compromesso.

Lo studio che si sta conducendo si avvale di lunghe serie di dati idro-meteorologici, ma nella pratica quotidiana risulta difficile disporre di una mole così consistente di osservazioni. Si vuole dunque investigare quali siano le potenzialità dell'analisi regionale proposta nel caso si disponga di serie di osservazioni di lunghezza inferiore. La limitatezza delle informazioni idrologiche comporta inevitabilmente un accorciamento anche dell'insieme di dati che vengono adottati per la calibrazione. Si simula questa condizione procedendo alla cross-calibrazione e successiva cross-validazione del modello adottando come periodo di calibrazione prima l'intervallo di 5 anni (01/01/1990-31/12/1994) e poi di soli 2 anni (01/01/1992-31/12/1993). In entrambi i casi, si adotta sempre di un intervallo di riscaldamento del modello pari a 305 giorni e le impostazioni generali della procedura rimangono le stesse adottate nel caso precedente.



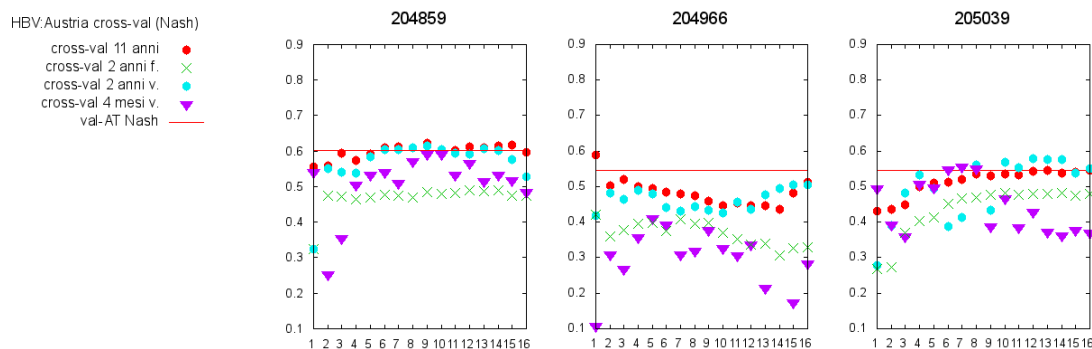
**Figura II- 10:** Prestazioni, in termini di efficienza di Nash, della cross-validazione del modello HBV, calibrato su periodi di lunghezza differente (11 anni in rosso, 5 anni in azzurro e 2 anni in verde).

Qualunque sia la lunghezza dell'intervallo adottato per la calibrazione della procedura regionale, se ne verificano le prestazioni avvalendosi unicamente delle osservazioni di portata registrate sul periodo 1976-1986 e 1998-2008. In Figura II- 10 si riportano i risultati espressi in termini di efficienza di Nash-Sutcliffe e si rimanda all'appendice (Figura A 1) per la visualizzazione delle efficienze calcolate sulle portate log-trasformate. Si nota come, mediamente, riducendo la lunghezza del periodo di calibrazione da 11 a 5 anni la procedura regionale riduca le potenzialità di previsione. Se si adottano serie idrologiche di lunghezza maggiori per la calibrazione di un modello, è molto probabile che queste contengano condizioni idrologiche estremamente varie ed è dunque possibile mettere a punto un modello che riesce a simulare meglio anche in fase di previsione situazioni idrologiche di diverso tipo (*Merz et al., 2009*). Un'ulteriore riduzione della lunghezza del periodo di calibrazione a 2 anni non comporta una conseguente significativa perdita di capacità predittiva.

### **3.2 Non contemporaneità delle osservazioni di portata su bacini differenti**

Di sovente le stazioni idrometriche che ricadono all'interno di un'area di ampia estensione non sono entrate in funzione tutte nello stesso momento; di conseguenza non solo ci si può trovare nella condizione di disporre di osservazioni di portata di limitata lunghezza, ma al contempo queste possono riferirsi a intervalli temporali differenti.

Si mima questa situazione adottando periodi di calibrazione di limitata lunghezza (pari a 2 anni e a 4 mesi soltanto) non più fissi per tutti i bacini oggetto di studio, ma variabili di volta in volta. Quando, ad esempio, si considera un periodo di calibrazione di 2 mesi e si procede alla cross-calibrazione regionale sul bacino target 205039 si sceglie di ipotizzare per il primo bacino donatore un periodo di calibrazione che si estende sull'intervallo 1992-1993; il secondo donatore viene invece calibrato su un periodo leggermente differente, il 1993-1994, così come per il terzo donatore si considera il periodo 1994-1995. Sostanzialmente a ciascun donatore compete un periodo di calibrazione sfasato rispetto a quello adottato per gli altri, ma scelto comunque in modo tale da adottare unicamente i dati che appartengono alla medesima finestra temporale che era stata scelta per le calibrazioni delle analisi precedenti. Lo stesso accade anche quando si considera un periodo di calibrazione di soli 4 mesi per ciascun donatore. Agendo in questo modo è possibile verificare la bontà della procedura adottata su dati indipendenti osservati nel periodo 1976-1986 e 1998-2008. Data la limitatezza del numero di osservazioni considerate nel caso della calibrazione su 4 mesi, si è scelto di portare il modello a regime sui 2 mesi precedenti ai 4 mesi effettivi che vengono utilizzati in calibrazione.



**Figura II- 11:** efficienze di Nash-Sutcliffe registrate in cross-validazione; si adottando periodi di calibrazione di lunghezza differente, sempre costanti per tutti i bacini donatori (cross-val 11 anni e cross-val 2 anni f.) oppure variabili (cross-val 2 anni v. e cross-val 4 mesi v.).

Per valutare le prestazioni dell'analisi regionale che si riferisce a periodi di calibrazione differenti di bacino in bacino, si è scelto di paragonare i risultati con quanto ottenuto dalla cross-validazione effettuata su periodi fissi. In Figura II- 11 si riporta a titolo di esempio, il caso della cross-validazione di soli tre bacini target mentre i dettagli dell'analisi condotta sull'intera area di studio sono riportati in appendice (Figura A 2). Il modello calibrato riferendosi ad informazioni provenienti su ciascun bacino da periodi totalmente differenti si dimostra di buona qualità, almeno per ciò che riguarda l'utilizzo di un periodo di calibrazione pari a 2 anni; si nota infatti come, mediamente, si raggiungano risultati molto migliori rispetto a quelli ai quali si perviene nel caso della calibrazione effettuata ricorrendo all'utilizzo degli stessi 2 anni per tutti i bacini donatori. In buona parte dei casi si raggiungono le efficienze tipiche della calibrazione effettuata su 11 anni. Adottando intervalli temporali variabili al variare del bacino donatore è possibile includere informazioni estremamente preziose per la calibrazione del modello. Come si è visto nel caso precedente, se si considera un intervallo temporale limitato è difficile riuscire a calibrare il modello affinché ben si adatti alle svariate condizioni idrologiche che si possono verificare in un bacino; considerando invece informazioni provenienti da periodi e da aree differenti il modello viene calibrato in condizioni idrologiche estremamente diverse e, di conseguenza, riesce a prevedere con maggiore facilità la risposta del bacino a sollecitazioni di vario genere.

E' invece eccessivamente ridotto un periodo di soli 4 mesi per ottenere simulazioni di buona qualità.

Sin qui si è verificata l'applicabilità della procedura di regionalizzazione proposta; si è visto cioè come, in media, per poter simulare la risposta idrologica di bacini non strumentati, sia possibile ricorrere all'utilizzo di un modello afflussi-deflussi il cui insieme di parametri derivi dalla calibrazione del modello su un numero consistente di bacini ad esso prossimi. In generale poi si osserva una limitata variazione delle prestazioni del modello all'aumentare del numero di bacini considerati per la calibrazione, indice del fatto che i bacini che appartengono a questa regione sono sostanzialmente molto simili.

Si è inoltre sperimentato che le informazioni provenienti da diversi bacini, se anche si riferiscono a intervalli temporali di limitata estensione e, soprattutto, non sono contemporanei,



sono estremamente utili per mettere a punto un modello in grado di ben riprodurre condizioni idrologiche di vario genere.

La verifica dell'applicabilità della medesima procedura di cross-calibrazione regionale anche in un contesto caratterizzato da condizioni morfo-climatiche completamente differenti e da serie di osservazioni di portata estremamente più corte viene affrontata nel prossimo capitolo.

## 4 Cross-calibrazione ed incertezza nei dati

Dall'analisi condotta per comprendere quale sia l'influenza del periodo di calibrazione sulla procedura di cross-validazione regionale si è visto come la regione dell'Innviertel sia costituita da 19 bacini il cui comportamento è sostanzialmente omogeneo. In un contesto del genere risulta difficoltoso individuare un trend, un'indicazione cioè di quale sia effettivamente il giusto numero di bacini donatori da adottare per ottenere una buona stima della risposta di un bacino idrologico non strumentato. Per limitare l'incertezza dovuta all'omogeneità dei bacini oggetto di studio si propongono due soluzioni:

- si considera un'area di studio completamente differente, quella cioè costituita dai bacini che ricadono nella regione "Emilia" il cui comportamento idrologico non presenta lo stesso grado di omogeneità,
- si limita l'attenzione unicamente ai bacini, di entrambe le aree di studio, che non sono annidati l'uno all'interno dell'altro.

### 4.1 Modelli idrologici e incertezza nei dati

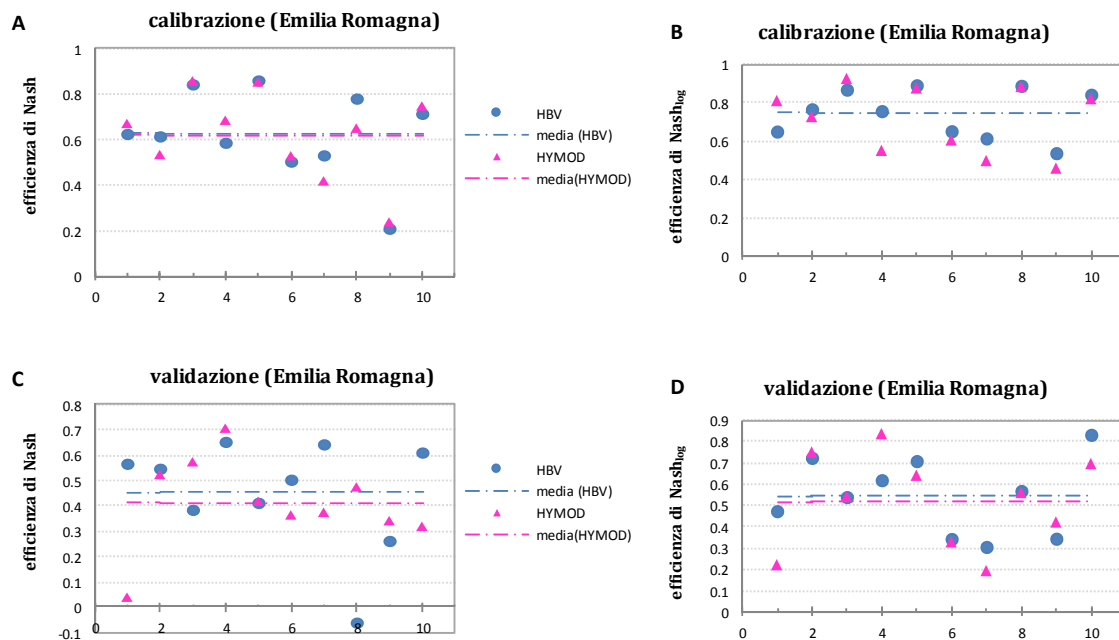
L'analisi di cross-calibrazione è in linea di principio realizzabile su qualunque regione e la si può sviluppare adottando qualsivoglia modello afflussi-deflussi. Rivolgendo l'attenzione alla regione "Emilia", regione per la quale si dispone di una base dati più limitata rispetto a quella studiata precedentemente, si vuole in questa sezione mettere a confronto le prestazioni di due modelli idrologici concettuali differenti per struttura e ampiamente adottati e descritti in letteratura; HYMOD è un modello a 5 soli parametri, dunque estremamente parsimonioso, mentre HBV necessita la calibrazione di almeno 13 parametri. A parità di prestazioni è bene propendere per la scelta del modello più semplice (*Bergstrom, 1991*) così da limitare l'incertezza associata ai parametri (Montanari, 2011); si è già visto come HBV sia più adatto alla rappresentazione dei fenomeni che avvengono nell'area dell'Innviertel e si vuole ora verificare cosa accade nella regione "Emilia".

I bacini che ricadono all'interno della regione "Emilia" sono stati inizialmente analizzati mediante una calibrazione locale e validazione di entrambi i modelli HBV e HYMOD per indagare quale sia più performante e dunque anche maggiormente adatto alle applicazioni regionali. Disponendo unicamente di 7 anni di osservazioni giornaliere di portata si è deciso di limitare il periodo di calibrazione a 2 anni (dal 01/01/2006 al 31/12/2007), di verificare i modelli sul restante periodo (dal 01/01/2003 al 31/12/2005 e dal 01/01/2008 al 31/12/2009) e di adottare un periodo di riscaldamento per il modello pari a 300 giorni. I risultati dell'analisi preliminare, riportati in Figura II- 12 mostrano come in fase di calibrazione i modelli portino mediamente ai medesimi risultati, essendo in grado di riprodurre sostanzialmente allo stesso modo sia le magre che i picchi di piena.

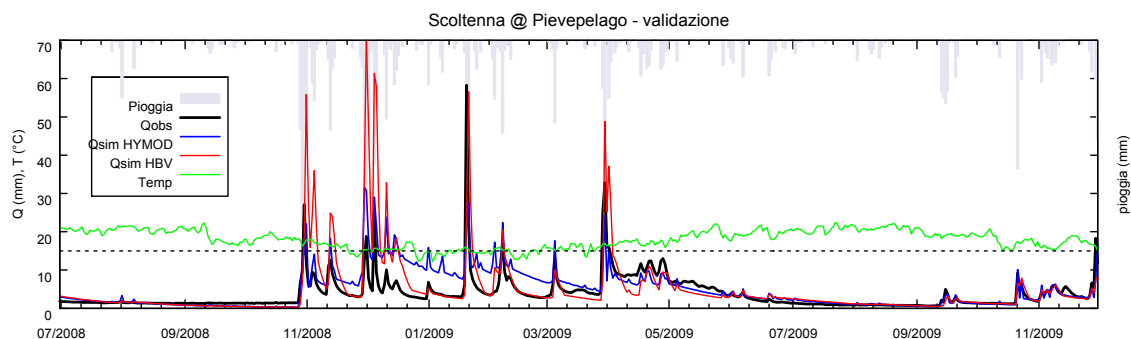
In fase di validazione le efficienze di Nash-Sutcliffe valutate sia sulle portate non trasformate che sulle portate log-trasformate, subiscono un leggero calo, attestandosi su valori leggermente maggiori nel caso si adotti il modello HBV, ma questa unica considerazione non giustifica la

scelta di un modello molto più oneroso. Si va dunque a verificare quali siano le ripercussioni, in termini di riproduzione dell'idrogramma, sui bacini della regione "Emilia" e si nota che mediamente in fase di calibrazione entrambi i modelli simulano parimenti il comportamento del bacino (si veda il riquadro in basso della Figura II- 14). Per quanto riguarda il comportamento in validazione dei due modelli si verifica che il ramo discendente dell'idrogramma di piena e le magre vengono meglio descritte, mediamente, dal modello HBV (ci si riferisca al riquadro in alto della Figura II- 14 e alla

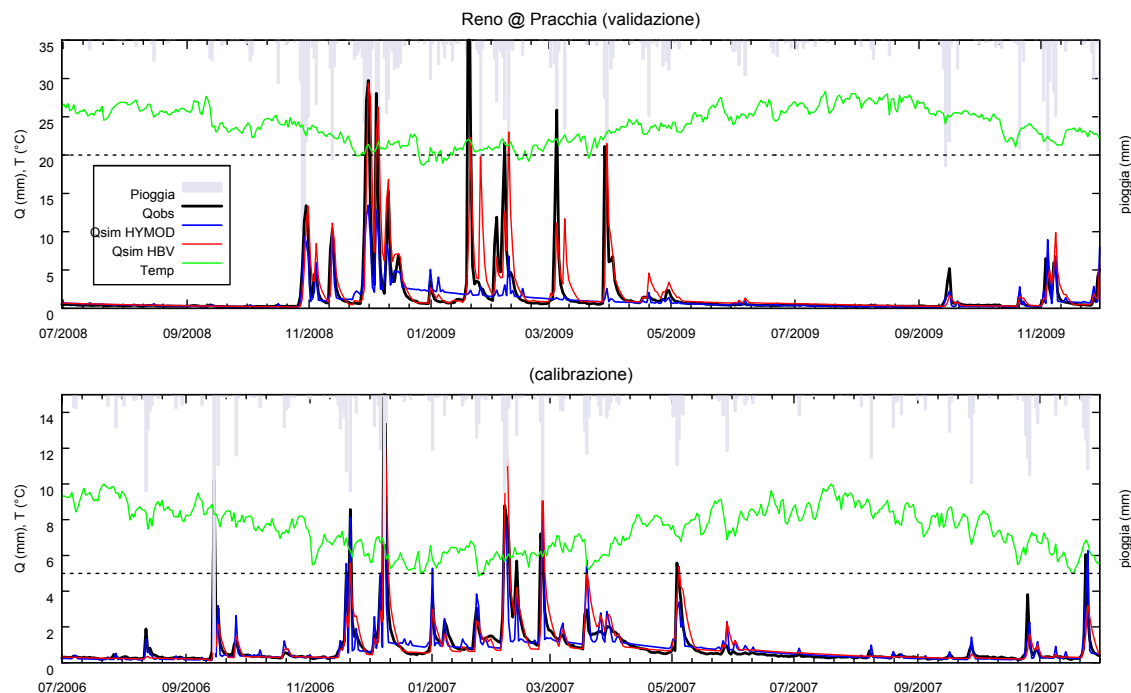
Figura II- 13).



**Figura II- 12:** Indici di bontà dei modelli HBV e HYMOD calibrati (A e B) e validati (C e D) nel senso classico sui 10 bacini emiliano-romagnoli considerati nell'analisi.



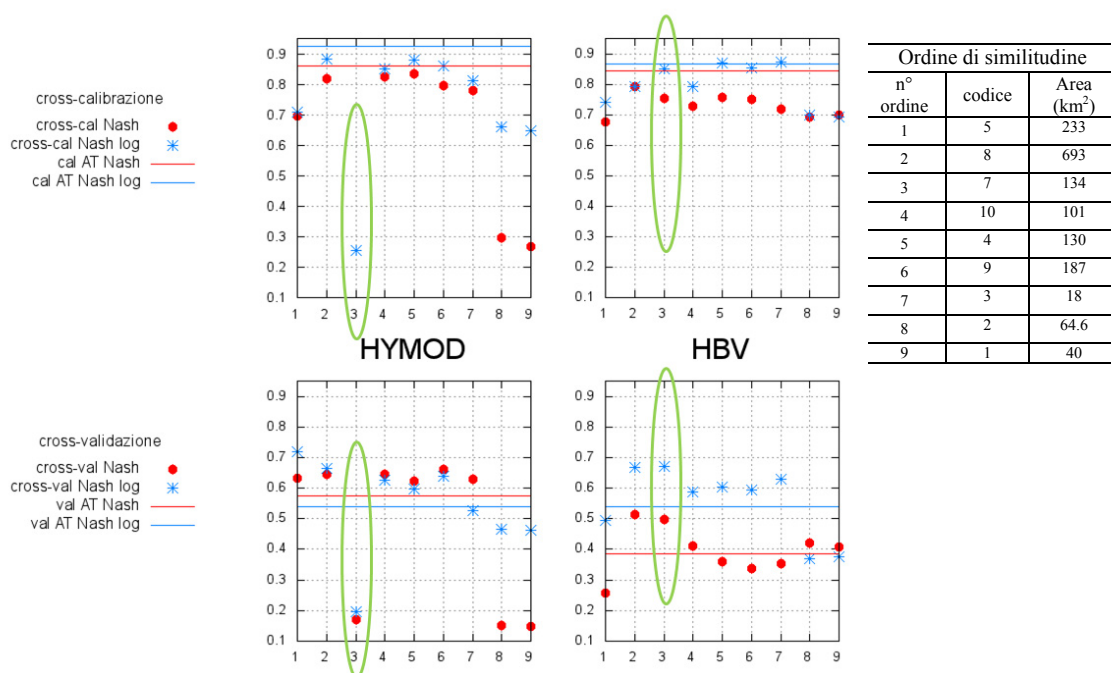
**Figura II- 13:** Simulazioni idrologiche realizzate con i modelli HBV e HYMOD sul bacino del torrente Scoltenna. (ci si focalizza su una parte del periodo di validazione).



**Figura II- 14:** Simulazioni idrologiche realizzate con i modelli HBV e HYMOD sul bacino del fiume Reno chiuso a Pracchia. In basso una porzione delle simulazioni effettuate per calibrare i modelli; in alto il riferimento è ad una porzione del periodo di validazione.

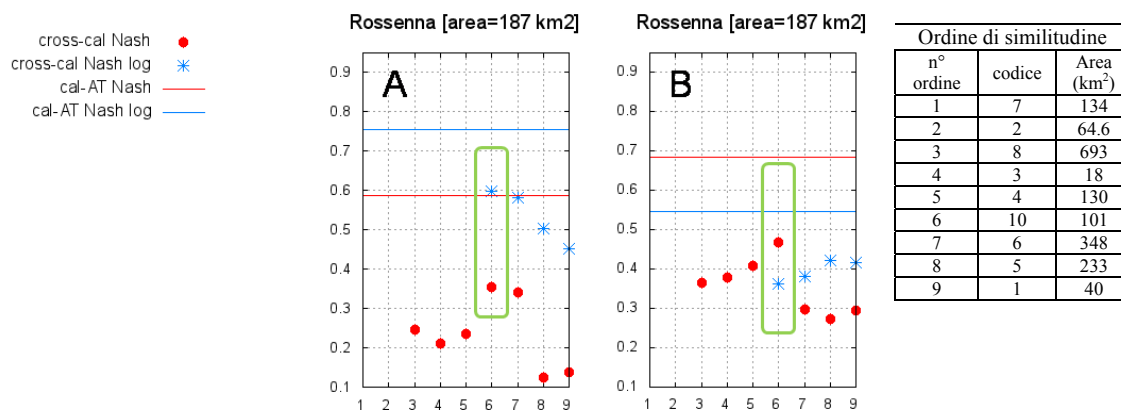
Si procede dunque alla cross-calibrazione di entrambi i modelli (HBV e HYMOD) applicati alla medesima area; si utilizzano le stesse impostazioni adottate per la calibrazione e validazione condotta in modo classico, effettuando dunque la cross-calibrazione sull'intervallo temporale 2006-2007 e la cross-validazione sui dati restanti. Ancora una volta l'ordine secondo il quale vengono considerati i bacini donatori per ciascun target viene stabilito in base alla vicinanza spaziale tra i bacini misurata secondo Ghosh. Si nota che i due modelli portano a risultati differenti, nel senso che se applicati al medesimo bacino target nella maggior parte dei casi non è possibile riconoscere il medesimo schema nella stima della bontà della procedura. E' possibile però riconoscere, così come si era notato nel caso dell'applicazione della metodologia alla regione dell'Innviertel, che mediamente adottando un buon numero di bacini donatori si riesce a determinare un set di parametri che, trasferito direttamente al bacino considerato non strumentato, porta ad una simulazione idrologica di buona qualità su quest'ultimo.

Dal confronto accurato dei risultati ottenuti con entrambi i modelli si è notato che il modello HBV, che prevede più del doppio di parametri rispetto ad HYMOD e che in fase di calibrazione e validazione classica su ciascun bacino appartenente alla regione "Emilia" ha presentato potenzialità del tutto paragonabili a quelle del modello più parsimonioso, si è rivelato estremamente interessante. Si è notato infatti che il modello HBV è più adatto a simulare situazioni nelle quali l'incertezza nei dati è significativa.



**Figura II- 15:** esito della procedura di cross-calibrazione (in alto) e cross-validazione (in basso) effettuate sul bacino del fiume Secchia chiuso a ponte Cavola adottando il modello HYMOD (a sinistra) e il modello HBV (a destra).

Si vedano a tal proposito i risultati della procedura di cross-calibrazione e di cross-validazione considerando come bacino non strumentato il bacino del fiume Secchia chiuso a ponte Cavola (Figura II- 15); i primi due donatori che vengono presi in considerazione sono i bacini del Secchia chiusi rispettivamente a Gatta (quindi a monte di ponte Cavola) e a Lugo (cioè a valle della sezione considerata come target). La cross-calibrazione e la cross-validazione porta dunque a buoni risultati, dal momento che si stanno considerando due bacini che sicuramente sono idrologicamente simili al target in questione essendo annidati in lui. Nel momento in cui si aggiunge un terzo donatore, il bacino del torrente Dolo, la procedura regionale ha un brusco calo di efficienza nel caso si consideri il modello HYMOD, mentre adottando il modello HBV la diminuzione di bontà della procedura è percettibile in validazione, ma non è altrettanto onerosa. Ogni qualvolta che, scegliendo il modello HYMOD per realizzare la procedura di cross-validazione, si aggiunge un donatore che non è idrologicamente sufficientemente simile al target oggetto di studio, i risultati presentano un brusco calo. L'esempio appena riportato è utile anche per sottolineare come la scelta di adottare la vicinanza geografica come criterio di similitudine sia discutibile in questo contesto.



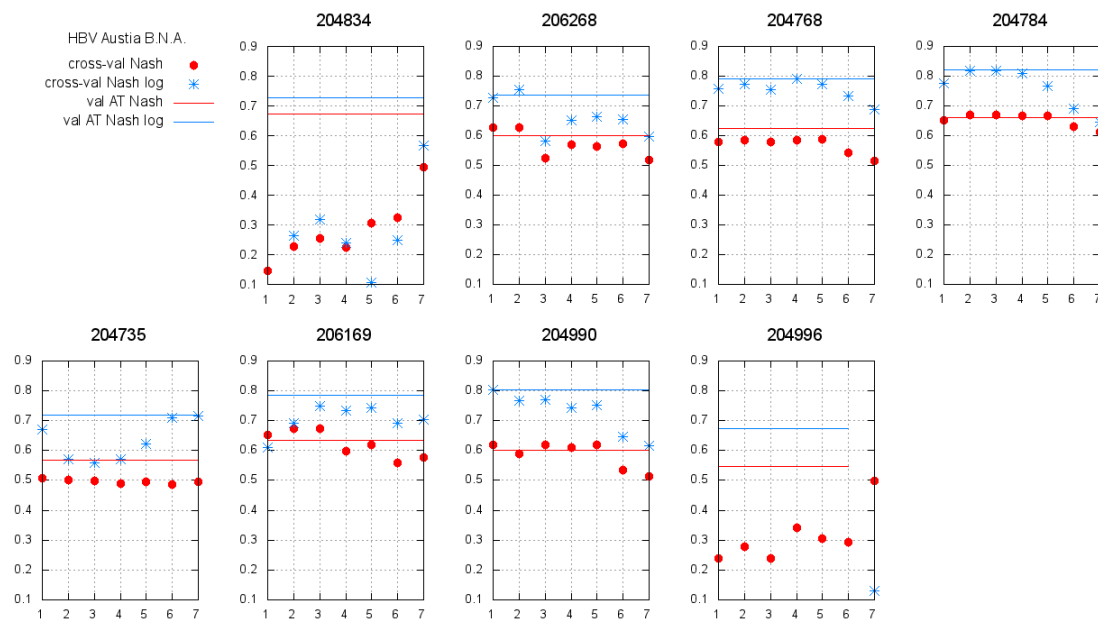
**Figura II- 16:** efficienza di Nash-Sutcliffe valutata sulle portate non trasformate (Nash) e su quelle log-trasformate (Nash log) nella fase di cross-calibrazione realizzata considerando il bacino del torrente Rossenna come target ed adottando il modello HBV (A) e HYMOD (B). Si evidenzia in verde l'adozione al sesto passo della procedura del bacino del torrente Tassobbio nel novero dei bacini donatori.

Si noti poi ad esempio quanto accade in fase di cross-calibrazione al bacino del torrente Rossenna (in Figura II- 16). I bacini donatori sono stati ordinati secondo distanze crescenti dal bacino del Rossenna (come riportato nella sezione di destra della medesima figura) e la procedura di cross-calibrazione prevede la calibrazione del modello inizialmente sul solo bacino del torrente Dolo, mentre al secondo passo la calibrazione dei parametri avviene considerando i bacini del torrente Dolo e del fiume Secchia chiuso a ponte Cavola; fino ad ora entrambi i modelli in cross-calibrazione portano a risultati talmente pessimi da non essere visualizzabili nel grafico delle efficienze positive. Solo quando all'insieme dei donatori viene aggiunto il bacino del torrente Tassobbio, al sesto passo, la procedura di cross-calibrazione porta a risultati accettabili. Ciò è vero in particolare se si utilizza il modello HBV (indicato con A in figura) e se si presta attenzione alla rappresentazione delle portate basse (si segua la rappresentazione dell'efficienza di Nash-Sutcliffe valutata sulle portate log trasformate in Figura II- 16). E' questa una riprova di quanto si era già notato in fase di calibrazione e validazione classica: con il modello HBV mediamente si rappresentano più fedelmente sia il ramo discendente dell'idrogramma di piena sia i bassi valori di portata. Al sesto passo della procedura si è effettivamente aggiunto agli altri bacini donatori, il bacino che più degli altri risulta idrologicamente simile al bacino target. Entrambi infatti non si estendono fino allo spartiacque appenninico e sono caratterizzati da pendenze medie inferiori rispetto agli altri bacini donatori adottati, oltre che da un regime climatico molto simile. E' questo dunque un segnale del fatto che la vicinanza geografica, nel caso della regione "Emilia" non è la migliore misura di similitudine idrologica. L'insieme completo dei risultati della cross-validazione realizzata sui bacini della regione "Emilia" adottando sia il modello HBV che il modello HYMOD viene riportato in appendice (Figura A 3 e Figura A 4).

## 4.2 Cross-calibrazione su bacini non annidati

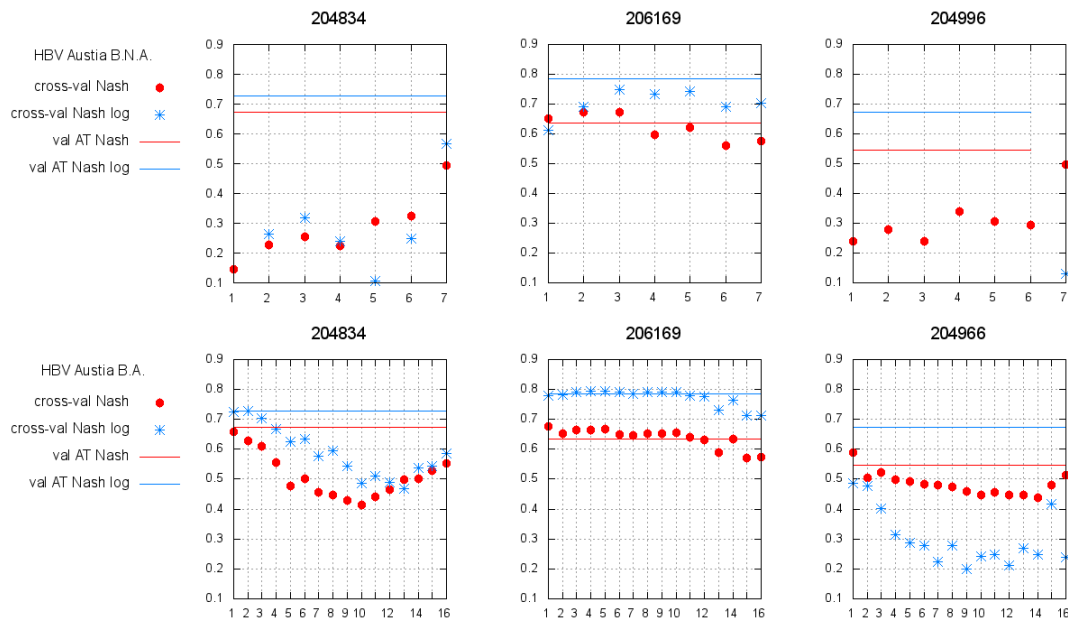
Il secondo metodo proposto per limitare l'incertezza dovuta all'omogeneità dei bacini oggetto di studio consiste nel realizzare le analisi di cross-calibrazione a scala regionale studiando unicamente i bacini che non sono annidati l'uno all'interno dell'altro. Si riduce quindi il numero di bacini disponibili per le analisi da 19 a 8 nel caso della regione dell'Innviertel e da 10 a 7 nel caso della regione "Emilia".

L'analisi di cross-calibrazione e di cross-validazione viene dunque ripetuta adottando un numero inferiore di bacini, ma mantenendo sempre le medesime impostazioni adottate fino a questo momento. Si presentano i risultati ottenuti analizzando sia i bacini che appartengono alla regione dell'Innviertel (sfruttando unicamente il modello HBV perché si è dimostrato più adatto) e di seguito quelli che riguardano i bacini che ricadono nella regione "Emilia".



**Figura II- 17:** Cross-validazione condotta adottando il modello HBV sui bacini della regione Innviertel non annidati (B.N.A.).

Anche per il caso di bacini austriaci non annidati non è possibile individuare un modello che si ripete inalterato su tutti i bacini target; non è al momento possibile dunque stabilire nettamente quale sia il numero (o il tipo) di bacini donatori da considerare contemporaneamente per pervenire alla migliore simulazione del comportamento idrologico di un bacino non strumentato che si trova nell'area dell'Innviertel. Esistono due casi in particolare che si discostano nettamente dal comportamento medio e che vengono analizzati di seguito nel dettaglio.



**Figura II- 18:** Paragone tra cross-validazione condotta sulla regione dell’Innviertel considerando solo i bacini non annidati (B.N.A.) e la cross-validazione considerando tutti i bacini disponibili.

Bacino target: 204834		Bacino target: 204966		Bacino target: 206169	
Codice donatore	B.N.A.	Codice donatore	B.N.A.	Codice donatore	B.N.A.
204826		204958		205047	
204859		205039		205054	
206268	X	205021		206326	
206326		204974		204990	X
206169	X	204990	X	204784	X
205054		205047		205021	
205047		206169	X	206102	
204784	X	206102		204859	
204768	X	205054		206268	X
204990	X	204784	X	205039	
206102		206805		204834	X
205021		206326		204974	
204974		204735	X	204826	
205039		204768	X	204768	X
206805		204859		204958	
204958		206268	X	204966	X
204735	X	204834	X	206805	
204966	X	204826		204735	X

**Tabella II- 4:** Elenco dei bacini donatori ordinati secondo distanza di Ghosh crescente dal bacino target. Nella cross-calibrazione effettuata su bacini non annidati si studiano unicamente i donatori identificati come B.N.A.

I risultati della cross-validazione condotta sul bacino 204834 (in Figura II- 18) considerando nell’analisi unicamente i bacini non annidati portano a tendenze completamente differenti rispetto a quelle verificatesi nel caso della cross-validazione condotta adottando tutti i bacini presenti nell’area. In Tabella II- 4 viene riportato l’ordine seguito per i bacini donatori considerati nel caso di cross-validazione sui soli bacini non annidati (identificati dalla colonna B.N.A.) e su tutti i bacini. Si consideri il caso di cross-validazione condotta adottando tutti i bacini della regione; il primo donatore per il bacino target 204834 risulta il bacino annidato identificato sul medesimo corso idrico e chiuso poco a monte rispetto al target. Il secondo donatore è il bacino annidato, identificato questa volta da una sezione di chiusura posta più a valle del bacino target. Guardando il riquadro in basso della Figura II- 18 si vede come le



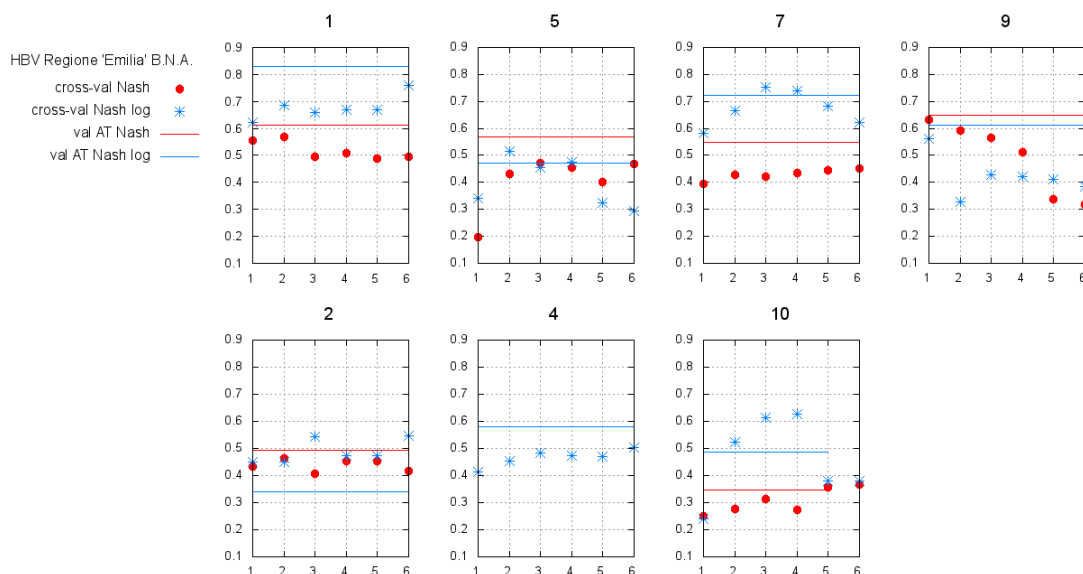
informazioni provenienti da questi due bacini siano estremamente importanti per ottenere una buona rappresentazione di quanto accada nel bacino 204834 considerato come non strumentato. Questi due bacini non vengono più considerati nel caso di cross-validazione condotta su bacini non annidati. Il primo donatore nell'analisi di bacini non annidati è invece il 206268, un bacino che si trova nelle immediate vicinanze del bacino target ma le cui informazioni non sono sufficientemente idonee alla calibrazione del set di parametri da adottare sul bacino target. Il modello realizzato adottando il set di parametri derivante dalla calibrazione del bacino 206268 non è infatti in grado di simulare quanto accade all'interno del bacino target, come si vede dal primo passo della cross-validazione rappresentata nel riquadro in alto della Figura II- 18. Solo nel momento in cui si considerano contemporaneamente tutti i bacini donatori è possibile stimare un insieme di parametri che, trasferito sul bacino target, porta a risultati accettabili. Si nota dunque quanto, per questo bacino, siano preziose le informazioni derivanti dai bacini annidati chiusi immediatamente a monte e a valle rispetto ad esso.

Si perviene alle medesime conclusioni osservando la tendenza riscontrata nell'andamento della bontà della cross-validazione condotta sul bacino 206966; nel caso si adottino unicamente i bacini non annidati le efficienze riportate in Figura II- 18 sono estremamente inferiori a quelle riscontrate nell'analisi condotta utilizzando tutti i bacini presenti nell'area. In questo caso infatti anche i bacini ad esso più prossimi non sono in grado di rappresentarne il comportamento idrologico; inoltre è doveroso sottolineare che questo bacino si colloca praticamente ai bordi di tutta l'area di studio, dunque i donatori considerati nell'analisi "non annidata" sono caratterizzati da condizioni differenti. Il primo donatore considerato nel caso di cross-validazione su bacini non annidati è infatti un bacino che non si trova nelle sue immediate vicinanze e che quindi è caratterizzato già da condizioni geo-morfo-climatiche non compatibili con esso.

Il bacino 206169 presenta invece, per il caso annidato e per quello non annidato, comportamento solo leggermente differente. Il primo bacino donatore utilizzato nel caso di cross-validazione condotta con bacini non annidati (il 4° donatore in assoluto) è un bacino che, pur non essendo annidato, si trova nelle immediate vicinanze e che ha dunque caratteristiche ad esso molto simili. Il modello però realizzato adottando unicamente i bacini del caso non annidato non è così performante come quello realizzato ricorrendo alle informazioni derivanti da tutti i bacini della zona. Da ciò si può desumere che esistono importanti similitudini anche tra bacini che non risultano annidati e che appartengono alla medesima area; le informazioni però più significative provengono dai bacini strumentati che si collocano immediatamente a monte e immediatamente a valle della sezione considerata non strumentata. Per questo, mediamente, la cross-validazione condotta utilizzando come bacini donatori i bacini situati a monte e a valle del bacino target porta a risultati giudicati molto buoni e, la continuazione del processo di cross-validazione aggiungendo ai bacini donatori anche bacini non annidati che presentano caratteristiche simili al bacino target, non porta ad uno scostamento significativo dei risultati.

Merz & Blöschl (2004) con uno studio condotto su 308 bacini austriaci pervengono ad un risultato analogo, sostenendo che la miglior tecnica di regionalizzazione dei parametri, tra le otto differenti da loro studiate, consiste nell'adottare la media tra i parametri calibrati localmente sui due bacini (annidati) che si trovano immediatamente a monte e a valle del bacino non strumentato.

L'analisi effettuata sui soli bacini non annidati della regione "Emilia" porta a risultati in qualche modo simili a quelli riscontrati sulla regione dell'Innviertel. Anche in questo caso non è possibile individuare un andamento tipico per tutti i bacini (si veda la Figura II- 19), ma se si paragonano i risultati con quelli ottenuti considerando tutti i bacini dell'area è possibile ottenere preziose informazioni sulla similitudine tra i bacini della zona. Si è scelto di adottare il modello HBV per procedere con questa ulteriore analisi perché, come si è sottolineato in precedenza, è più adatto rispetto ad HYMOD alla descrizione di situazioni nelle quali si riscontra incertezza nei dati.



**Figura II- 19:** Cross-validazione condotta adottando il modello HBV sui bacini della regione "Emilia" non annidati (B.N.A.).

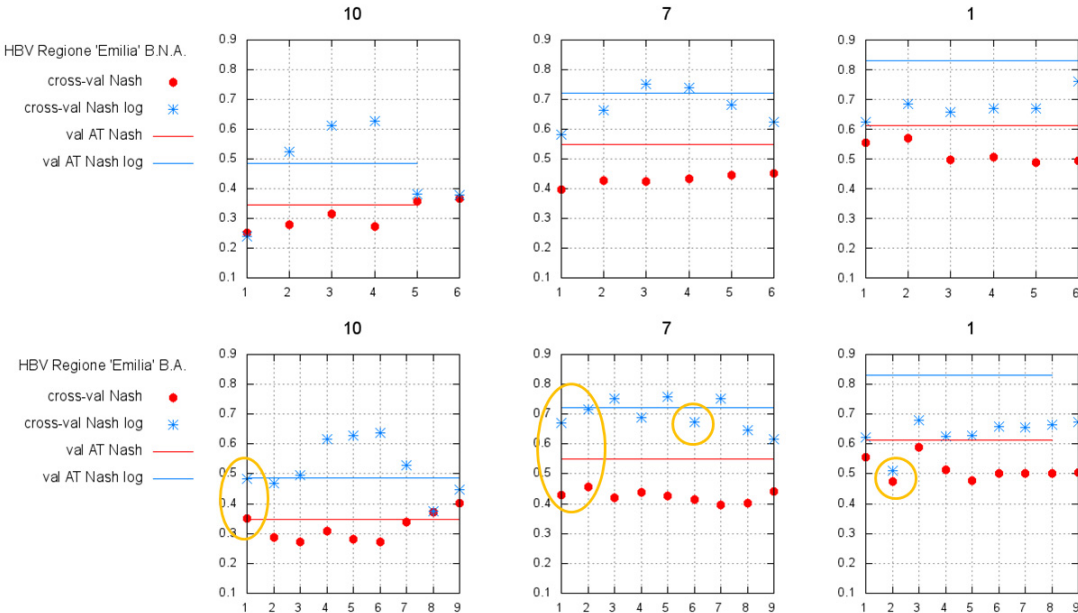
Una conferma del fatto che bacini simili si possono trovare nelle immediate vicinanze del bacino non strumentato viene dal bacino 10 (il bacino del torrente Tassobbio) per il quale si rappresentano in Figura II- 20 i risultati ottenuti in cross-validazione adottando unicamente l'insieme di bacini non annidati (sopra) ed anche quelli presentati precedentemente per l'analisi condotta su tutti i bacini dell'area (sotto). In Tabella II- 5 è poi riportato l'ordine secondo il quale si sono considerati i bacini donatori e si nota come il primo donatore del caso completo, evidenziato in figura, sia un bacino che non viene considerato successivamente nell'analisi sui bacini non annidati. Non si tratta però di un bacino che comprende il bacino target. Si sta infatti considerando il bacino del torrente Tassobbio come bacino non strumentato e il suo primo donatore nel caso di cross-validazione su tutti i bacini dell'area è il bacino del fiume Secchia chiuso a ponte Cavola. Il mancato utilizzo delle informazioni che provengono da quest'ultimo, nel caso di cross-validazione sui bacini non annidati, comporta un drastico peggioramento della qualità del processo.

Se si considera invece il bacino numero 7 (torrente Dolo a ponte Dolo) come non strumentato e si paragonano i risultati delle analisi condotte con tutti i bacini dell'area e solo con i bacini non annidati si nota l'importanza delle informazioni provenienti dal bacino nel quale è annidato (8, Secchia a Lugo). Venendo a mancare questo bacino nel caso di cross-

validazione sui bacini non annidati, il peggioramento delle prestazioni del modello è significativo.

Si notano infine risvolti positivi dovuti all'aver escluso dall'analisi alcuni bacini. Il bacino donatore 6 (Secchia a ponte Cavola) provoca un peggioramento delle condizioni se viene considerato nel processo di cross-validazione del bacino target 7 (Dolo a ponte Dolo); una volta rimosso dal novero dei donatori ecco che l'andamento del grafico delle efficienze tipiche della cross-validazione condotta sul bacino 7 assume un andamento più regolare. Lo stesso accade sul bacino target 1 (Reno a Pracchia): eliminando il bacino donatore 3 (Acquicciola a Fiumalbo) dalle analisi si ottiene un notevole beneficio.

Si può dunque concludere che, per quanto riguarda la regione "Emilia", si riconosce l'importanza della vicinanza spaziale come indicatore di similitudine idrologica tra i bacini, ma si ritiene che non sia in grado di spiegare completamente le dinamiche di similitudine.



**Figura II- 20:** Regione "Emilia", paragone tra cross-validazione effettuata considerando tutti i bacini presenti nella regione (in basso) e quella realizzata considerando solo i bacini non annidati (B.N.A. in alto).

		Bacino target: 10		Bacino target: 7		Bacino target: 1	
ordine	codice	B.N.A.		codice	B.N.A.	codice	B.N.A.
1	6			8		2	X
2	5	X		6		3	
3	8			5	X	4	X
4	7	X		4	X	9	X
5	9	X		9	X	7	X
6	4	X		3		8	
7	3			10	X	6	
8	2	X		2	X	5	X
9	1	X		1	X	10	X

**Tabella II- 5:** Regione "Emilia", elenco dei bacini donatori ordinati secondo distanza di Ghosh crescente dal bacino target. B.N.A indica i bacini considerati per le analisi sui soli bacini non annidati.

## 5 Considerazioni conclusive della II parte

Le analisi di regionalizzazione dei parametri adottate solitamente in letteratura si rivelano estremamente lunghe e talvolta portano a risultati non soddisfacenti; la ricerca infatti di relazioni che correlino ciascun parametro del modello afflussi-deflussi alle caratteristiche fisiche proprie dei bacini appartenenti alla regione di studio è influenzata da un alto grado di incertezza. Si sceglie invece di adottare una metodologia alternativa che punti alla regionalizzazione non di ciascun parametro, ma alla calibrazione regionale dell'intero insieme di parametri e per farlo si adottano le informazioni provenienti dai bacini strumentati più idrologicamente simili al sito di interesse. In particolare si assume l'ipotesi di invarianza dei parametri sui bacini donatori sui quali viene calibrato contemporaneamente il set di parametri da trasferire, senza alcuna ulteriore modifica, al bacino non strumentato. Si è voluto con questo studio comprendere quale sia il giusto compromesso tra lo svantaggio derivante dall'assumere parametri omogenei sui bacini donatori e il vantaggio di adottare informazioni provenienti da un numero sempre crescente di bacini.

Lo studio, condotto su due differenti aree (la regione dell'Innviertel in Austria, e la regione "Emilia") e adottando due modelli idrologici differenti per struttura e soprattutto per complessità, evidenzia i vantaggi derivanti dall'assunzione di parametri provenienti dalla calibrazione contemporanea del modello su un certo numero di bacini donatori; pur non essendo in grado, al momento, di stabilire quale sia il numero e il tipo ottimale di bacini donatori sui quali realizzare la cross calibrazione, si è convinti della bontà dell'analisi regionale proposta. Si è infatti provato, su entrambe le aree di studio, che le efficienze proprie del modello regionale in validazione sono paragonabili con le efficienze del modello che venga calibrato e validato localmente sugli stessi bacini target considerati.

Si riconosce però che la misura di similitudine adottata in questo studio, la vicinanza spaziale valutata secondo Ghosh, non è in grado di spiegare completamente le dinamiche di similitudine tra i bacini. Questo è soprattutto vero per i bacini che appartengono alla regione "Emilia". Le informazioni che provengono dai bacini annidati che sono identificabili tramite sezioni poste a monte e a valle del sito di interesse risultano estremamente preziose, sia per i bacini austriaci che per quelli appenninici; i bacini invece contigui possono risultare molto simili, e dunque apportano benefici al processo di cross-calibrazione (è il caso del bacino del fiume Secchia chiuso a ponte Cavola che risulta molto importante per la corretta calibrazione del torrente Dolo), ma possono presentare anche caratteristiche totalmente differenti provocando un peggioramento della procedura. Si è dunque convinti della necessità di identificare una misura di similitudine tra bacini che tenga in considerazione contemporaneamente la prossimità geografica ed anche caratteristiche differenti.

Contestualmente si è prestata particolare attenzione a due aspetti: da un lato si è voluto indagare se la scelta di un modello idrologico in particolare influenzi l'intera procedura di

calibrazione regionale, mentre d'altro canto si è studiata la rilevanza del periodo adottato per la cross-calibrazione.

Il modello più complicato e più oneroso, HBV, si è rivelato in realtà più idoneo, rispetto ad HYMOD, alla simulazione regionale in caso di incertezza nei dati; si è visto infatti come HBV risenta in misura minore della presenza all'interno dell'insieme di donatori, di bacini che non risultano sufficientemente simili al bacino non strumentato. Per quanto riguarda invece la scelta del periodo di calibrazione da adottare, si è visto, come ci si aspettava, che un periodo di calibrazione lungo (compatibilmente con quelli che sono le serie di osservazioni di portata disponibili nelle stazioni idrometriche presenti nell'area di studio) sia la scelta da perseguire ogni qualvolta ce ne sia la possibilità. Qualora invece si abbiano a disposizione unicamente serie di osservazioni di lunghezza limitata nel tempo, come si è visto nel caso della regione dell'Innviertel adottando un periodo di calibrazione pari a 2 anni, è estremamente utile ricorrere alla calibrazione regionale sfruttando informazioni che provengono da diversi bacini e che sono state registrate in periodi non concomitanti. I risultati ottenuti in fase di previsione calibrando il modello su un numero alto di bacini ed adottando informazioni non contemporanee ha infatti portato a prestazioni molto superiori a quelle ottenute dalla calibrazione del modello sullo stesso numero di osservazioni, ma concomitanti; si è visto infatti come i risultati ottenuti siano paragonabili a quelli propri di una calibrazione realizzata su un periodo di 11 anni.



## Stima dell'evapotraspirazione con il Radiation Method.

Per ogni bacino idrografico, la stima del valore dell'evapotraspirazione avviene secondo i seguenti passi, come suggerito da Todini (1996): (1) stima della temperatura media mensile di lungo periodo in ogni stazione di misura; (2) stima del gradiente termico; (3) stima della temperatura media mensile di lungo periodo per l'intero bacino (si tratta in questo caso del valore di temperatura media areale e non puntuale); (4) stima dell'evapotraspirazione potenziale media mensile calcolata secondo la formula di Thornthwaite; (5) stima dei coefficienti di regressione  $\alpha$  e  $\beta$  per ogni bacino.

- (1) Per ogni stazione meteo, per la quale si disponga di un adeguato numero di anni di dati osservati (Todini (1996) consiglia di adottare solo le stazioni per le quali esistano serie di dati di lunghezza minima pari a 10 anni) si stimano i valori di temperatura media mensile di lungo periodo.
- (2) Se il bacino oggetto di studio è caratterizzato da un'ampia escursione topografica e le stazioni termometriche sono poste a quota sensibilmente variabile, allora la stima del gradiente termico è di rilevante importanza. Posizione geografica ed momento dell'anno sono fattori che influenzano il gradiente termico, al punto che disponendo di un cospicuo numero di stazioni omogeneamente distribuite su un bacino sarebbe possibile determinare un valore differente giorno per giorno. Ci si limita in questo caso a determinare, e a sfruttare per la stima dell'evapotraspirazione, un valore medio.
- (3) I valori di temperatura media mensile di lungo periodo ottenuti per ogni stazione di misura vengono ragguagliati ad una medesima quota, la quota media del bacino, per mezzo del gradiente termico. Grazie alla tecnica dei poligoni di Thiessen si determina la serie di temperature medie mensili areali corrispondenti all'area di interesse.
- (4) Poiché il ruolo giocato dall'evapotraspirazione è determinante in termini di riduzione dell'umidità del suolo nel tempo, è indispensabile che la stima dell'evapotraspirazione media nel lungo periodo prodotta dal modello scelto sia accurata. I valori mensili di evapotraspirazione potenziale da utilizzarsi per la calibrazione dei parametri del *radiation method* sono determinati ricorrendo alla formula di Thornthwait and Mather (1955):

$$ET_{pm}(k) = 16 \cdot \frac{n(k)}{30} \frac{N(k)}{12} \left[ 10 \cdot \frac{\bar{T}_m(k)}{\vartheta} \right]^{\gamma} \quad (\text{A.1})$$

nella quale

- $ET_{pm}(k)$  è l'evapotraspirazione potenziale media mensile (mm day<sup>-1</sup>) del mese  $k$ ,

- $\bar{T}_m(k)$  (°C) è la temperatura mensile media di lungo periodo riferita al mese  $k$ ,
  - $N(k)$  (-) è la media mensile, riferita al mese  $k$ , del numero massimo di ore di sole giornaliere tabulato in funzione della latitudine (si veda Tabella A- 2),
  - $n(k)$  (-) è il numero di giorni presenti nel mese  $k$ ,
  - $\theta = \sum_{k=1}^{12} \left[ \frac{\bar{T}_m(k)}{5} \right]^{1.514}$  è l'indice termico, dato dalla somma degli indici mensili,
  - $\gamma = 0.49239 + 1792 \cdot 10^{-5} \theta - 771 \cdot 10^{-7} \theta^2 + 675 \cdot 10^{-9} \theta^3$
- (5) La stima dell'evapotraspirazione potenziale mensile è valutata in AffDef come linearmente dipendente dalla temperatura, secondo la:

$$ET_{pm}(k) = \alpha + \beta \cdot N(k) \cdot W_{TA}(k) \cdot \bar{T}_m(k) \quad (\text{A.2})$$

nella quale:  $\alpha$  (mm) e  $\beta$  (mm°C<sup>-1</sup>) sono i coefficienti di regressione che necessitano di essere calibrati.  $W_{TA}(k)$  (-) è il fattore di compensazione, funzione della temperatura media mensile e dell'altitudine, determinabile grazie ai valori tabulati in Doorembos et al. (1975) e qui riportati in Tabella A- 1 che sono ben approssimabili attraverso la parabola:

$$W_{TA}(k) = A + B \cdot \bar{T}(k) + C \cdot \bar{T}(k)^2 \quad (\text{A.3})$$

dove  $\bar{T}$  è la temperatura media di lungo periodo (°C) riportata in Tabella A- 1 e A, B, C sono i tre parametri di regressione. Todini (1996) dimostra che le stime fornite dall'equazione (A.2) approssimano bene i valori di evapotraspirazione ottenuti tramite la (A.1), fornendo così una corretta calibrazione dei parametri  $\alpha$  e  $\beta$ .

Poiché poi l'equazione (A.2) è lineare nei confronti della temperatura, una volta stimati i parametri  $\alpha$  e  $\beta$ , l'equazione seguente può essere adottata con passo di calcolo giornaliero, più appropriato per il caso scelto:

$$ET_{pd}(t) = \{\dot{\alpha} + \dot{\beta} \cdot N(k) \cdot W_{TA}(k) \cdot T_m(t)\} / n(k) \quad (\text{A.4})$$

nella quale  $ET_{pd}(t)$  è l'evapotraspirazione stimata per il passo temporale scelto (mm d<sup>-1</sup>),  $\dot{\alpha}$  e  $\dot{\beta}$  sono i valori stimati per i due parametri del modello e  $T_m(t)$  è il dato osservato di temperatura media giornaliera (°C).

W <sub>ta</sub> Quota [m]	Temperatura [°C]																			
	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40
0	0.43	0.46	0.49	0.52	0.55	0.58	0.61	0.64	0.66	0.71	0.73	0.75	0.77	0.78	0.85	0.82	0.83	0.84	0.85	0.86
500	0.45	0.48	0.51	0.54	0.57	0.60	0.62	0.65	0.70	0.72	0.74	0.76	0.78	0.79	0.81	0.82	0.84	0.85	0.86	0.87
1000	0.46	0.49	0.52	0.55	0.58	0.61	0.66	0.69	0.71	0.73	0.75	0.77	0.79	0.80	0.82	0.83	0.85	0.86	0.87	0.88
2000	0.49	0.52	0.55	0.58	0.61	0.64	0.66	0.69	0.71	0.73	0.75	0.77	0.79	0.81	0.82	0.84	0.85	0.86	0.87	0.88
3000	0.52	0.55	0.58	0.61	0.64	0.66	0.69	0.71	0.73	0.75	0.77	0.79	0.81	0.82	0.84	0.85	0.86	0.88	0.88	0.89
4000	0.53	0.58	0.61	0.64	0.66	0.69	0.71	0.73	0.76	0.78	0.79	0.81	0.83	0.84	0.85	0.86	0.88	0.89	0.90	0.90

**Tabella A- 1:** Valori del fattore di compensazione riportati in corrispondenza di differenti quote

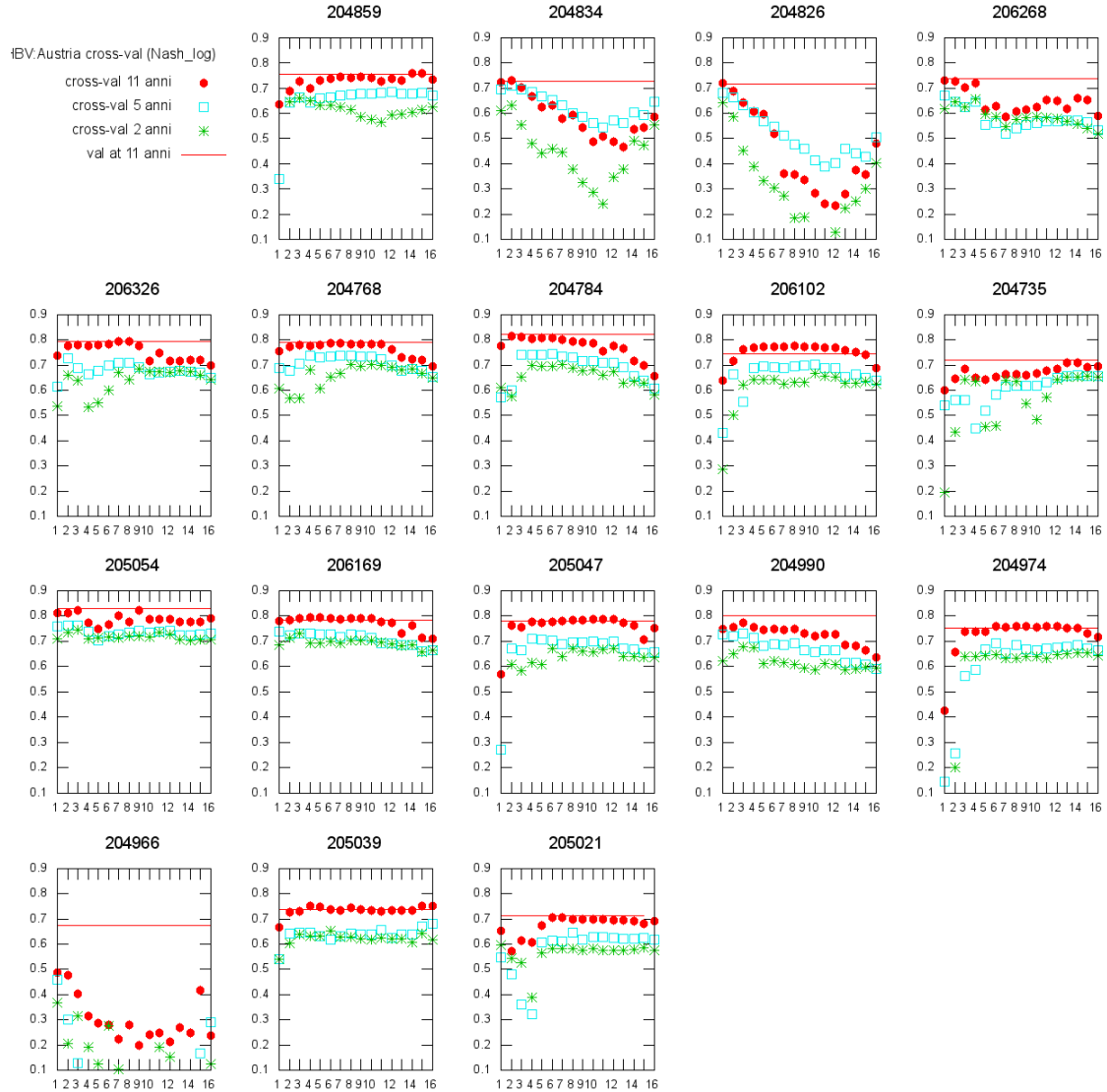


Lat. Nord	Gen	Feb	Mar	Apr	Mag	Giu	Lug	Ago	Set	Ott	Nov	Dic
Lat. Sud	Lug	Ago	Set	Ott	Nov	Dic	Gen	Feb	Mar	Apr	Mag	Giu
<b>50</b>	8.5	10.1	11.8	13.8	15.4	16.3	15.9	14.5	12.7	10.8	9.1	8.1
<b>48</b>	8.8	10.2	11.8	13.6	15.2	16.0	15.6	14.3	12.6	10.9	9.3	8.3
<b>46</b>	9.1	10.4	11.9	13.5	14.9	15.7	15.4	14.2	12.6	10.9	9.5	8.7
<b>44</b>	9.3	10.5	11.9	13.4	14.7	15.4	15.2	14.0	12.6	11.0	9.7	8.9
<b>42</b>	9.4	10.6	11.9	13.4	14.6	15.2	14.9	13.9	12.6	11.1	9.8	9.1
<b>40</b>	9.6	10.7	11.9	13.3	14.4	15.0	14.7	13.7	12.5	11.2	10.0	9.3
<b>35</b>	10.1	11.0	11.9	13.1	14.0	14.5	14.3	13.5	12.4	11.3	10.3	9.8
<b>30</b>	10.4	11.1	12.0	12.9	13.6	14.0	13.9	13.2	12.4	11.5	10.6	10.2
<b>25</b>	10.7	11.3	12.0	12.7	13.3	13.7	13.5	13.0	12.3	11.6	10.9	10.6
<b>20</b>	11.0	11.5	12.0	12.6	13.1	13.3	13.2	12.8	12.3	11.7	11.2	10.9
<b>15</b>	11.3	11.6	12.0	12.5	12.8	13.0	12.9	12.6	12.2	11.8	11.4	11.2
<b>10</b>	11.6	11.8	12.0	12.3	12.6	12.7	12.6	12.4	12.1	11.8	11.6	11.5
<b>5</b>	11.8	11.9	12.0	12.2	12.3	12.4	12.3	12.3	12.1	12.0	11.9	11.8
<b>0</b>	12.1	12.1	12.1	12.1	12.1	12.1	12.1	12.1	12.1	12.1	12.1	12.1

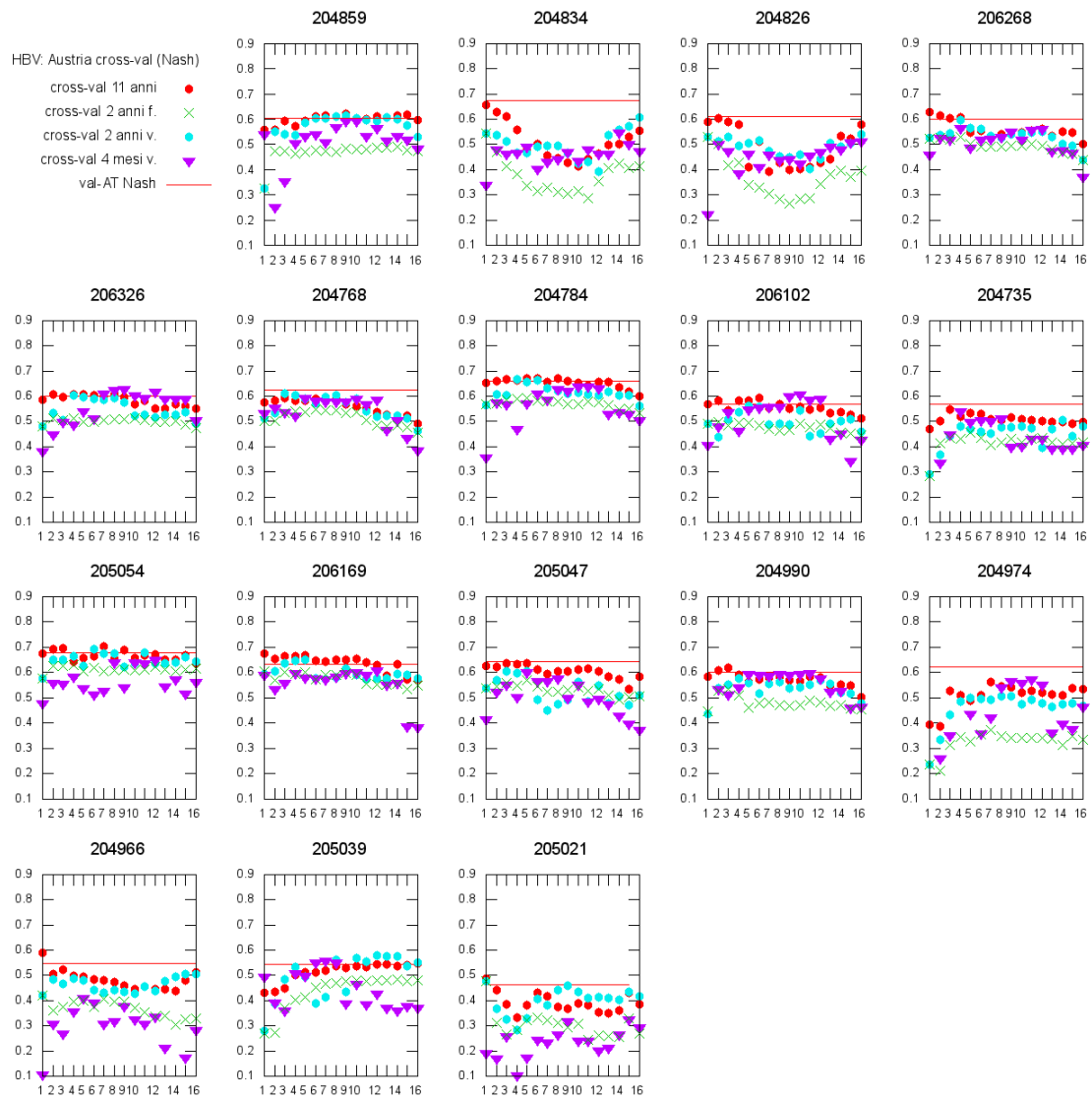
**Tabella A- 2:** Durata giornaliera media del numero massimo di ore di sole in funzione del mese e della latitudine (riprodotta dalla Tabella 3 di 'Crop Water Requirement', FAO Irrigation and Drainage Paper 24).

# Analisi Regionale

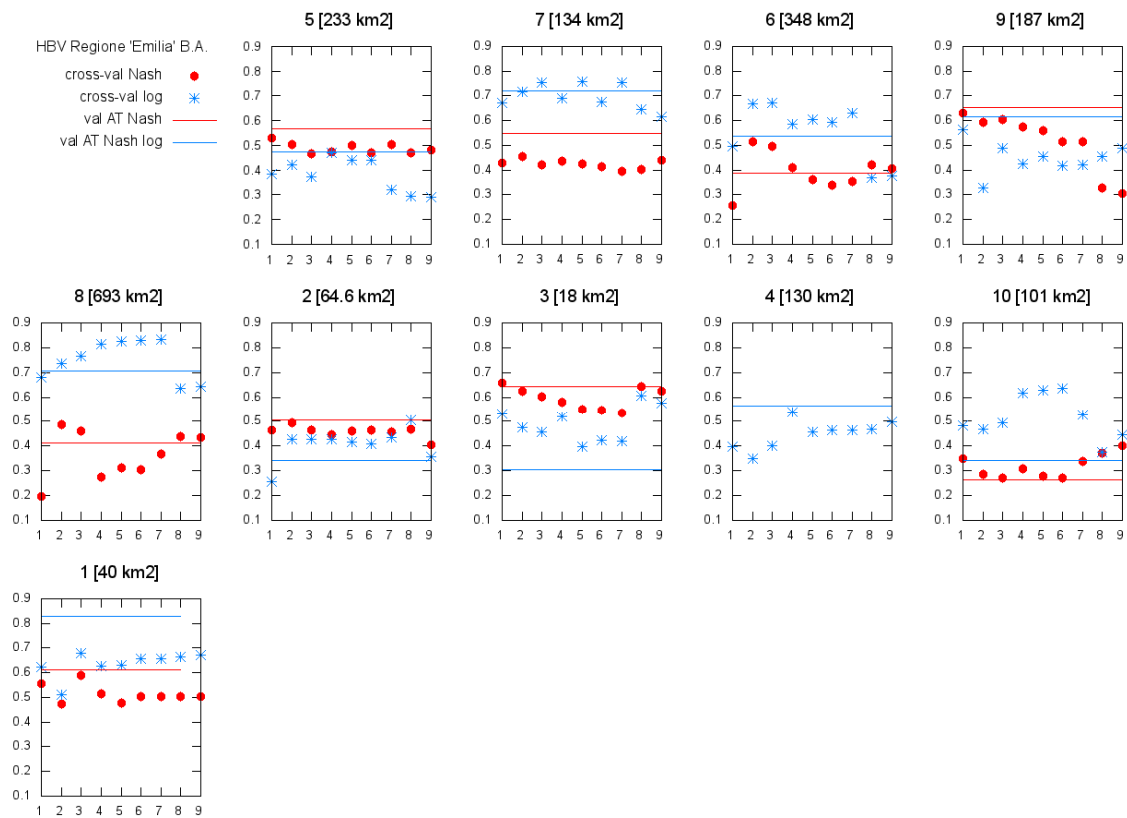
Si propongono in queste pagine alcuni grafici ad integrazione di quanto riportato nella Parte II di questo lavoro di tesi.



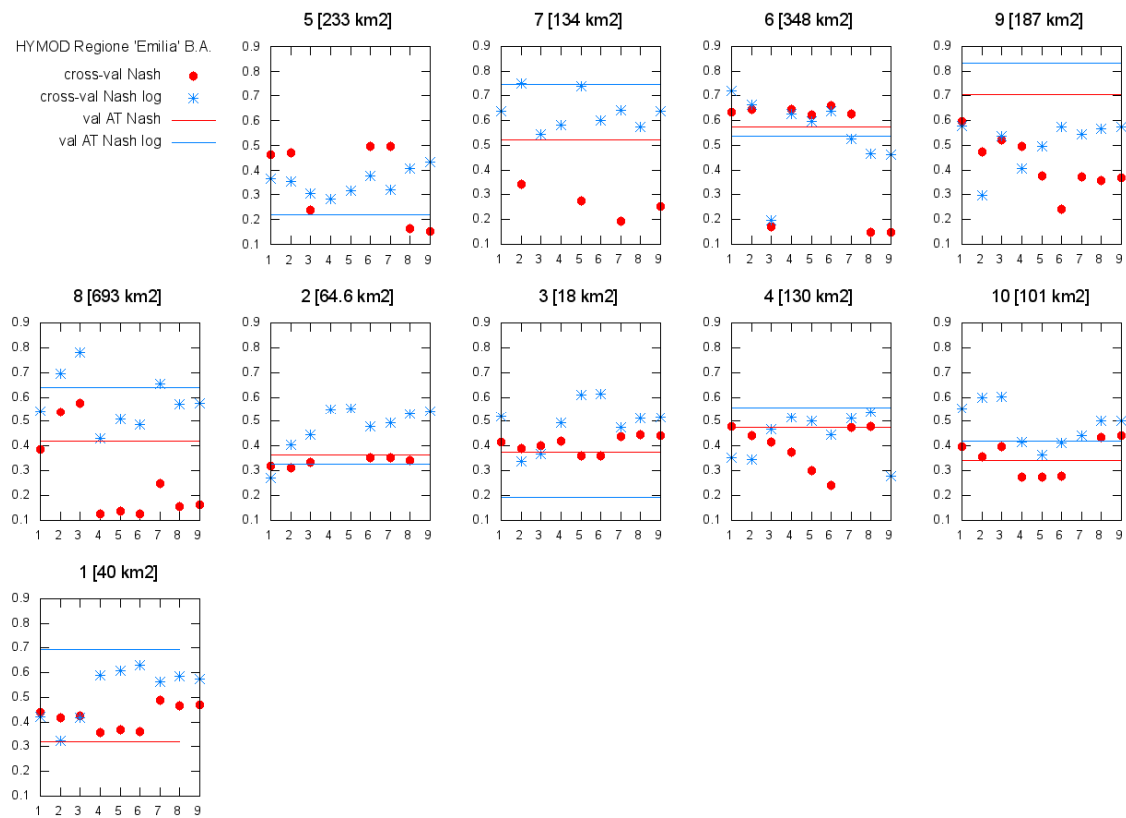
**Figura A 1:** Prestazioni, in termini di efficienza di Nash-Sutcliffe calcolata sulle osservazioni di portata log-trasformate, della cross-validazione del modello HBV, calibrato su periodi di lunghezza differente (11 anni in rosso, 5 anni in azzurro e 2 anni in verde).



**Figura A 2:** Non contemporaneità delle osservazioni di portata su bacini differenti: efficienze di Nash-Sutcliffe registrate in cross-validazione. si adottando periodi di calibrazione di lunghezza differente, sempre costanti per tutti i bacini donatori (cross-val 11 anni e cross-val 2 anni f.) oppure variabili (cross-val 2 anni v. e cross-val 4 mesi v.).



**Figura A 3:** Risultati della cross-validazione (cross-val) condotta su tutti i bacini della regione “Emilia”; il modello HBV è stato calibrato sul periodo 2006-2007 sia con una calibrazione e validazione classica (val AT) sia mediante la procedura di cross-calibrazione. Si valuta la bontà del modello in termini di efficienza di Nash-Sutcliffe valutata sulle portate non trasformate (Nash) e sulle portate log-trasformate (Nash log).



**Figura A 4:** Risultati della cross-validazione (cross-val) condotta su tutti i bacini della regione “Emilia”; il modello HYMOD è stato calibrato sul periodo 2006-2007 sia con una calibrazione e validazione classica (val AT) sia mediante la procedura di cross-calibrazione. Si valuta la bontà del modello in termini di efficienza di Nash-Sutcliffe valutata sulle portate non trasformate (Nash) e sulle portate log-trasformate (Nash log).

# Bibliografia:

- Anderson, E., Snow Accumulation and Ablation Model – SNOW-17, 2006  
([http://www.nws.noaa.gov/oh/hrl/nwsrfs/users\\_manual/part2/\\_pdf/22snow17.pdf](http://www.nws.noaa.gov/oh/hrl/nwsrfs/users_manual/part2/_pdf/22snow17.pdf))
- Arnold, J.G., Srinivasan, R., Mattiah, R.S., Williams, J.R., Large area hydrologic modeling and assessment. Part I: model development., Journal of the American Water Resources Association, 1998, 34, 73-89.
- Bates, B. and Campbell, E.P., A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling, Water Resources Research, 2001, 37, 4, 937-947.
- Bergstrom and Sandberg, Simulation of groundwater response by conceptual models – Three case studies, Nordic Hydrology, 1983, 14, 71-84
- Bergstrom S., “Chapter 13: The HBV model.” in *Computer models of watershed hydrology*, V. P. Singh, ed., Water Resources Publications, Littleton, Colo., 1995.
- Bergström, B. and Carlsson, B., River runoff to the Baltic Sea: 1950-1990. Ambio, 1994, 23, 280-287.
- Bergstrom, S., Development and application of a conceptual runoff model for Scandinavian catchments, Swedish Meteorological Institute, SMHI, Report no. RHO 7, Norrköping, 1976
- Bergstrom, S., Principles and Confidence in Hydrological Modelling, Nordic Hydrology, 1991, 22, 123-136
- Bergström, S., The HBV model - its structure and applications. SMHI Reports RH, No. 4, Norrköping, 1992
- Beven, K. Rainfall-runoff modeling – the primer. Chichester: Wiley, 2001
- Beven, K., A manifesto for the equifinality thesis, Journal of Hydrology, 2006, 320, 18-36.
- Beven, K. and Binley, A., The future of distributed models: model calibration and uncertainty prediction. Hydrological processes, 1992, 6, 279-298.
- Bloschl, G., Rainfall-runoff modelling of ungauged catchments, in *Encyclopedia of Hydrological Sciences*, John Wiley, Chichester, U.K., 2005.
- Bloschl, G., Raszler, C., Komma, J., A spatially distributed flash flood forecasting model, Environmental Modelling & Software, 2008, 23, 464-478.
- Boyle, D. P., H. V. Gupta, and S. Sorooshian, Toward improved calibration of hydrological models: Combining the strengths of manual and automatic methods, Water Resour. Res., 2000, 36, 3663–3674.
- Cunge, J.A., On the subject of a flood propagation computation method (Muskingum method), Journal of Hydraulic Research, 1969 7, 205-230

Doorenbos J. And Pruitt W. O., Guidelines for predicting crop water requirements, Irrigation and drainage paper 24, FAO 1975 (<http://www.fao.org/documents/en/detail/21139>)

Duan, Q., Sorooshian, S., and Gupta, V., Effective and efficient Global Optimizations for Conceptual Rainfall-Runoff models, Water Resources Research, 1992, 28(4), 1015-1031.

Ghosh, B., Random distances within a rectangle and between two rectangles., Bulletin of the Calcutta Mathematical Society, 1951

Gottschalk, L., Advances in observational hydrology-fields experiments and modeling. In: Takeuchi, K.,(Ed.), Proceedings of Workshop on the Prediction in Ungauged Basins (PUBs), 2002

Gottschalk, L., Interactive comment on “Catchments as space-time filters – a joint spatio-temporalgeostatistical analysis of runoff and precipitation” by J. O. Skøien and G. Blöschl, Hydrol. Earth Syst. Sci. Discuss., 2006, 3, S390–S393 ([www.hydrol-earth-syst-sci-discuss.net/3/S390/2006/](http://www.hydrol-earth-syst-sci-discuss.net/3/S390/2006/))

Gottschalk, L., Leblois, E., Skoien, J.O., Distance measures for hydrological data having a support., Journal of hydrology, 2011, 402, 415-421

Gupta, H., Sorooshian, S., Yapo, P., Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information., Water Resources Research. 1998, 34(4), 751-763. (doi:10.1029/97WR03495).

Harlin, J., Modelling the Hydrological Response of Extreme Floods in Sweden, Nordic Hydrology, 1992, 23, 227-244

Hock, R., Temperature index melt modeling in mountain areas., Journal of Hydrology, 2003, 282, 104-115.

Johansson, B., The relationship between catchment characteristics and the parameters of a conceptual runoff model - A study in the south of Sweden. Contribution to the Second International Conference on FRIEND, 1993, Braunschweig, IAHS Publication No. 221, 475-482.

Lindstrom, G., Johansson, B., Persson, M., Gardelin, M., Bergstrom, Development and test of the distributed HBV-96 hydrological model, Journal of Hydrology, 1997, 201, 272-288

Maidment D.R., Handbook of Hydrology, McGraw-Hill, Inc., New York, 1992

Mebane, W.R. & Sekhon, J.S., Genetic Optimization Using Derivatives: the rgenoud package for R., Polit. Anal., 1998, 7, 187-210

Merz, R. and Blöschl, G., Regionalisation of catchment model parameters., J. Hydrol., 2004, 287, 95-123

Merz, R., Parajka, J., Blöschl, G., Scale effects in conceptual hydrological modeling. Water Resources Research, 2009, 45, W09405 (doi:10.1029/2009WR007872).

Montanari, A. & Toth, E., Calibration of hydrological models in the spectral domain: an opportunity for scarcely gauged basins?, Water Resources Research, 2007, 43, W05434 (doi:10.1029/2006WR005184)

Montanari, A., What do we mean by “uncertainty”? The need for a consistent wording about uncertainty assessment in hydrology. *Hydrological Processes*, 2007, 21, 841-845

Moore, R.J., The probability-distributed principle and runoff production at point at basin scales., *Hydrological Sciences*, 1985, 30, 273-297

Nash, J. & Sutcliffe J., River flow forecasting through conceptual models, Part 1: a discussion of principles, *J. Hydrol.*, 1970, 10, 282-290

Nester, T., Kirnbauer, R., Gutknecht, D., Bloeschl, G., Climate and catchment controls on the performance of regional flood simulations, *Journal of Hydrology*, 2011, 402, 340–356

Oudin, L., V. Andréassian, C. Perrin, C. Michel, and N. Le Moine, Spatial proximity, physical similarity, regression and ungauged catchments: A comparison of regionalization approaches based on 913 French catchments, *Water Resour. Res.*, 2008, 44, W03413, (doi:10.1029/2007WR006240).

Parajka, J., Bloeschl, G., The value of MODIS snow cover data in validating and calibrating conceptual hydrologic models., *Journal of Hydrology*, 2008, 358, 240– 258.

Parajka, J., Bloeschl, G., Merz, R., Regional calibration of catchment models: potential for ungauged catchments., *Water Resources Research*, 2007a, 43, W06406, (doi:10.1029/2006WR005271)

Parajka, J., Merz, R. and Bloeschl, G., A comparison of regionalization methods for catchment model parameters. *Hydrol. Earth Syst. Sci.*, 2005, 9, 157-171

Parajka, J., Merz, R., Bloeschl, G., Uncertainty and multiple objective calibration in regional water balance modeling: case study in 320 Austrian catchments., *Hydrological Processes*, 2007, 21, 435-446

Pellicciotti, F., Brock, B., Strasser, U., Burlando, P., Funk, M., Corripio, J., An enhances temperature-index Glacier melt model including the shortwave radiation balance: development and testing for Haut Glacier d’Arolla, Switzerland., *Journal of Glaciology*, 2005, 51(175), 573-587.

Penman, H. L.. “Weather, plant and soil factors in hydrology.” *Weather*, 1961, 16, 207–219.

Perrin C., Michael, C., Andreassian, V., Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments., *Journal of Hydrology*, 2001, 242, 275-301.

Rango, A., “Chapter 14: The snowmelt runoff model (SRM)” in *Computer models of watershed hydrology*, V. P. Singh, ed., Water Resources Publications, Littleton, Colo., 1995, 443–476

R.E.R. (Agenzia Regionale Prevenzione e Ambiente Regione Emilia Romagna), *Annali Idrologici*, (2001-2010)

R.E.R. (Agenzia Regionale Prevenzione e Ambiente Regione Emilia Romagna), *Atlante idro-climatico dell’Emilia-Romagna 1961-2008*, 2010 (ISBN 88-87854-24-6).



Singh, V. P., “Chapter 1: Watershed modeling.” *Computer models of watershed hydrology*, V. P. Singh, ed., Water Resources Publications, Littleton, Colo., 1995, 1–22

Sivapalan M., Prediction in ungauged basins: a grand challenge for theoretical hydrology. *Hydrological Processes*, 2003, 17(15), 3163, 3170.

Sivapalan, M., et al., IAHS Decade on Predictions in Ungauged Basins (PUB), 2003-2012: Shaping an exciting future for the hydrological sciences, *Hydrological Sciences Journal*, 2003, 48(6), 857, 880, (DOI: 10.1623/hysj.48.6.857.51421)

Skoien, J.O. and Bloeschl, G., Spatiotemporal topological kriging of runoff time series, *Water Resources Research*, 2007, 43, W09419

Skoien, J.O., Pebesma, E.J., Bloeschl, G., rtop – an R package for interpolation of data with a non-point support., 2009 [Presented at EGU General assembly, (poster)]

Soil Conservation service, National Engineering Handbook, 1972. Section 4, Hydrology. U.S. Department of Agriculture, Washington DC

Sorooshian, S., and Gupta, V.K., Automatic calibration of conceptual rainfall runoff models: The question of parameter observability and uniqueness, *Water Resources Research*, 19(1), 251-259 (doi:10.1029/WR019i001p00260)

Thiessen, A. H, Precipitation averages for large areas., *Mon.Weather Rev.*, 1911, 39(7), 1082–1084

Thomthwaite, C.W. and Mather, J.R., The water balance. *Publications in Climatology*, 1955-8(1). Laboratory of Climatology, Centerton, NJ.

Todini, E., “The ARNO Rainfall-runoff model”, *Journal of Hydrology*, 1996, 175, 339-382

Todini, E., Rainfall-Runoff modeling – past, present and future. *Journal of Hydrology*, 1988, 100, 341-352.

Uhlenbrook, S., Seibert, J., Leibundgut, and A. Rodhe, Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure, *Hydrol. Sci., J.*, 1999, 44, 779-797.

Vandewiele, G.L., and Elias, A., Monthly water balance of ungauged catchments obtained by geographical regionalization., *Journal of Hydrology*, 1995, 170, 277-291.

Vrugt, J. A., W. Bouten, H. V. Gupta, and S. Sorooshian, Toward improved identifiability of hydrologic model parameters: The information content of experimental data, *Water Resour. Res.*, 2002, 38(12), 1312 (doi:10.1029/2001WR001118).

Vrugt, J.A., ter Braak, C. J. F., Clark, M. P., Hyman, J. M., Robinson, B. A., Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation, *Water Resources Research*, 2007, 44, W00B09 (doi:10.1029/2007WR006720).

Vrugt, J.A., ter Braak, C. J. F., Clark, M. P., Hyman, J. M. and Robinson, B. A., Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?, *Stoch Environ Res Risk Assess*, 2009 (doi:10.1007/s00477-008-0274-y).

Wagener, T., D. P. Boyle, M. J. Lees, H. S. Wheater, H. V. Gupta, and S. Sorooshian, A framework for development and application of hydrological models, Hydrol. Earth Syst. Sci., 2001, 5(1), 13– 26.

Wagener, T., Wheater HS., Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty., Journal of Hydrology, 2006, 320, 132-154.

Yadav, M., Wagener, T., Gupta, H., Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins., Advances in water resources, 2007, 30, 1756-1774.

[http://www.macauley.ac.uk/hydalp/private/demonstrator\\_v2.0/models/hbv.html](http://www.macauley.ac.uk/hydalp/private/demonstrator_v2.0/models/hbv.html) [HBV model]

# Elenco delle Figure:

## PARTE I:

Figura I- 1: Rappresentazione schematica della simulazione dei processi di accumulo e scioglimento nivale .....	9
Figura I- 2: Rappresentazione schematica dei fenomeni di interazione tra suolo, atmosfera e vegetazione riprodotti dal modello AFFDEF in ogni cella di discretizzazione del bacino.....	13
Figura I- 3: Schema di funzionamento del modello AFFDEF e dei parametri di calibrazione .....	16
Figura I- 4: Il problema dei minimi locali.....	22
Figura I- 5: A) funzione di densità di probabilità riferita ad un generico parametro, B) valori delle portate osservate, in nero, e simulate, in rosso, rappresentate tramite bande di incertezza al 90%. .....	24
Figura I- 6: Modello digitale del terreno (80 m.) dell'area di studio: il bacino del fiume Panaro. In mappa sono indicate anche le posizioni delle stazioni idrometriche (cerchi blu), termometriche (cerchi rossi) e pluviometriche (triangoli gialli). .....	33
Figura I- 7: Evoluzione della funzione log-verosimiglianza per ogni differente catena Markov (ad ogni colore corrisponde una MCMC diversa per un totale di 10 catene).....	38
Figura I- 8: Valori assunti dai parametri del modello AFFDEF all'interno di ogni differente catena Markov (ad ogni colore corrisponde una MCMC diversa per un totale di 10 catene). Le unità di misura sono quelle effettivamente adottate durante la calibrazione del modello.....	39
Figura I- 9: Risultato della calibrazione del modello AFFDEF sul bacino del torrente Leo presentati in forma di frequenza assoluta e rapportati all'intervallo di variabilità del parametro scelto a priori.....	40
Figura I- 10: Funzioni densità di probabilità a posteriori dei parametri di AFFDEF calibrato sul bacino del torrente Leo.....	41
Figura I- 11: Simulazioni idrologiche condotte sul bacino del torrente Leo: valori di portate osservate(in nero) e fasce di incertezza al 90% delle portate simulate. Si distingue il periodo di calibrazione (in rosso) da quello di validazione (in verde) .....	42
Figura I- 12: Funzioni di densità di probabilità a posteriori riferite ai parametri del modello AFFDEF calibrato sui bacini dei torrenti Leo, Scoltenna e del fiume Panaro. ....	43
Figura I- 13: registrazioni di temperatura media giornaliera della stazione di Doccia Fiumalbo (in blu), riferite allo 0 termometrico (linea nera tratteggiata) e valori di portata osservata (in nero) e simulata (in rosso) alla stazione di Pievepelago .....	45
Figura I- 14: Simulazioni idrologiche condotte sul bacino del torrente Scoltenna chiuso alla sezione di Pievepelago .....	46

Figura I- 15: Simulazioni idrologiche condotte sul bacino del fiume Panaro chiuso alla sezione di ponte Samone .....	47
Figura I- 16: distribuzioni di densità di probabilità di un parametro che ipoteticamente è stato calibrato su due differenti bacini.....	47
Figura I- 17: Consistenza della rete delle stazioni di misura di portata in Emilia Romagna al 31/12/2001 (in giallo), 31/12/2003 (in verde), 31/12/2010 (in rosso); ( <i>R.E.R, 2001, 2003, 2010</i> ). .....	50
Figura I- 18: il caso del torrente Scoltenna calibrato, seguendo le tre fasi, sul periodo 2003-2007.....	55
Figura I- 19: Il caso del torrente Scoltenna validato sull'anno 2006. ....	56
Figura I- 20: torrente Scoltenna (anno di validazione 2006): simulazioni idrologiche effettuate in accordo alla fase 1 e sfruttando il set di valori ottimali ottenuti per il torrente Leo. ....	58
 <b>PARTE II:</b>	
Figura II- 1: Regione dell'Innviertel, localizzata nella parte nord dell'Austria. (fonte: <i>Skoien and Bloeschl, 2007</i> ).....	65
Figura II- 2: bacini oggetto di studio presenti nell'area compresa tra i fiumi Enza, Secchia, Panaro e Reno; in blu sono indicate le stazioni idrometriche, in rosso le stazioni termometriche e in verde quelle pluviometriche.....	68
Figura II- 3: Prestazioni dei modello HBV e AFFDEF (nella versione concentrata) valutate in termini di efficienza di Nash-Sutcliffe.....	69
Figura II- 4: Struttura generale del modello HBV nella versione dello SMHI e pensato come applicato ad un sotto-bacino. (fonte: <a href="http://www.macaulay.ac.uk/hydalp/private/demonstrator_v2.0/models/hbv.html">http://www.macaulay.ac.uk/hydalp/private/demonstrator_v2.0/models/hbv.html</a> ).....	72
Figura II- 5: Schematizzazione del modulo responsabile della stima dell'umidità del suolo. FC, LP e BETA sono parametri empirici. (fonte: <i>Bergstrom, 1995</i> ).....	73
Figura II- 6: Schema di funzionamento del modello HYMOD .....	77
Figura II- 7: Indicatori di bontà dei modelli HBV e HYMOD riferiti alla fase di calibrazione (A e B) e validazione (C e D) effettuata nel senso classico sui soli 8 bacini austriaci non annidati. ....	81
Figura II- 8: Cross-calibrazione applicata alla regione dell'Innviertel adottando un periodo di calibrazione pari a 11 anni (01/01/1987-31/01/1997); risultati della calibrazione (locale, denominata cal-AT, e regionale, cross-cal) di 9 bacini espressi in termini di efficienza di Nash-Sutcliffe (Nash, in rosso) e di efficienza di Nash-Sutcliffe valutata sul logaritmo delle portate (Nash log, in blu).....	83
Figura II- 9: Cross-validazione realizzata sugli anni 1976-1986 e 1998-2008 considerando di volta in volta un bacino target differente. Si valuta la bontà della procedura regionale paragonando i risultati ottenuti in cross-validazione (punti) e attraverso la validazione classica del modello (linea) su ciascun bacino target in termini di efficienza di Nash calcolata sulle portate osservate (in rosso) e sulle portate log-trasformate (in blu).....	84
Figura II- 10: Prestazioni, in termini di efficienza di Nash, della cross-validazione del modello HBV, calibrato su periodi di lunghezza differente (11 anni in rosso, 5 anni in azzurro e 2 anni in verde).....	86
Figura II- 11: efficienze di Nash-Sutcliffe registrate in cross-validazione; si adottando periodi di calibrazione di lunghezza differente, sempre costanti per tutti i bacini donatori (cross-val 11 anni e cross-val 2 anni f.) oppure variabili (cross-val 2 anni v. e cross-val 4 mesi v.). ....	88

Figura II- 12: Indici di bontà dei modelli HBV e HYMOD calibrati (A e B) e validati (C e D) nel senso classico sui 10 bacini emiliano-romagnoli considerati nell'analisi.....	91
Figura II- 13: Simulazioni idrologiche realizzate con i modelli HBV e HYMOD sul bacino del torrente Scoltenna. (ci si focalizza su una parte del periodo di validazione).....	91
Figura II- 14: Simulazioni idrologiche realizzate con i modelli HBV e HYMOD sul bacino del fiume Reno chiuso a Pracchia. In basso una porzione delle simulazioni effettuate per calibrare i modelli; in alto il riferimento è ad una porzione del periodo di validazione. ....	92
Figura II- 15: esito della procedura di cross-calibrazione (in alto) e cross-validazione (in basso) effettuate sul bacino del fiume Secchia chiuso a ponte Cavola adottando il modello HYMOD (a sinistra) e il modello HBV (a destra). ....	93
Figura II- 16: efficienza di Nash-Sutcliffe valutata sulle portate non trasformate (Nash) e su quelle log-trasformate (Nash log) nella fase di cross-calibrazione realizzata considerando il bacino del torrente Rossenna come target ed adottando il modello HBV (A) e HYMOD (B). Si evidenzia in verde l'adozione al sesto passo della procedura del bacino del torrente Tassobbio nel novero dei bacini donatori. ....	94
Figura II- 17: Cross-validazione condotta adottando il modello HBV sui bacini della regione Innviertel non annidati (B.N.A.).....	95
Figura II- 18: Paragone tra cross-validazione condotta sulla regione dell'Innviertel considerando solo i bacini non annidati (B.N.A.) e la cross-validazione considerando tutti i bacini disponibili.....	96
Figura II- 19: Cross-validazione condotta adottando il modello HBV sui bacini della regione "Emilia" non annidati (B.N.A.).....	98
Figura II- 20: Regione "Emilia", paragone tra cross-validazione effettuata considerando tutti i bacini presenti nella regione (in basso) e quella realizzata considerando solo i bacini non annidati (B.N.A. in alto). ....	99
<b>Appendice:</b>	
Figura A 1: Prestazioni, in termini di efficienza di Nash-Sutcliffe calcolata sulle osservazioni di portata log-trasformate, della cross-validazione del modello HBV, calibrato su periodi di lunghezza differente (11 anni in rosso, 5 anni in azzurro e 2 anni in verde). ....	106
Figura A 2: Non contemporaneità delle osservazioni di portata su bacini differenti: efficienze di Nash-Sutcliffe registrate in cross-validazione. si adottando periodi di calibrazione di lunghezza differente, sempre costanti per tutti i bacini donatori (cross-val 11 anni e cross-val 2 anni f.) oppure variabili (cross-val 2 anni v. e cross-val 4 mesi v.). ....	107
Figura A 3: Risultati della cross-validazione (cross-val) condotta su tutti i bacini della regione "Emilia"; il modello HBV è stato calibrato sul periodo 2006-2007 sia con una calibrazione e validazione classica (val AT) sia mediante la procedura di cross-calibrazione. Si valuta la bontà del modello in termini di efficienza di Nash-Sutcliffe valutata sulle portate non trasformate (Nash) e sulle portate log-trasformate (Nash log). ....	108
Figura A 4: Risultati della cross-validazione (cross-val) condotta su tutti i bacini della regione "Emilia"; il modello HYMOD è stato calibrato sul periodo 2006-2007 sia con una calibrazione e validazione classica (val AT) sia mediante la procedura di cross-calibrazione. Si valuta la bontà del modello in termini di efficienza di Nash-Sutcliffe valutata sulle portate non trasformate (Nash) e sulle portate log-trasformate (Nash log). ....	109

## Elenco delle Tabelle:

Tabella I- 1: Elenco dei parametri del modello AFFDEF che richiedono la calibrazione .....	18
Tabella I- 2: Stazioni pluviometriche e termometriche utili allo studio. La numerazione riportata si riferisce alla collocazione delle stazioni riportate in Figura I- 6. ....	35
Tabella I- 3: intervalli di validità dei parametri. ( <b><i>Mf</i></b> espresso in $\log_{10}(\text{mm}/^{\circ}\text{C/s})$ ; <b><i>Hs</i></b> e <b><i>K</i></b> espressi in $\log_{10}(\text{s})$ come utilizzati dal modello) .....	37
Tabella I- 4: Elenco delle stazioni di misura di portata in funzione al 31/12/2001. ( <i>R.E.R., 2001</i> ).....	50
Tabella I- 5: Valori dell'efficienza di Nash calcolati in calibrazione (C) e in validazione (V) per le simulazioni idrologiche realizzate seguendo l'ipotesi di omogeneità dei parametri nello spazio. ....	54
Tabella I- 6: Valori assunti dai parametri del modello AFFDEF calibrato sul bacino del torrente Scoltenna sul periodo (2003-2005, 2007-2008) secondo le tre diverse fasi previste dall'analisi e valori degli stessi parametri ottenuti dalla calibrazione del modello sul bacino del fiume Leo (allo scopo di verificare il comportamento del bacino dello Scoltenna come se fosse non strumentato). ....	57
Tabella II- 1: Indicazione delle sezioni di chiusura considerate all'interno della regione di studio austriaca.....	66
Tabella II- 2: Indicazione delle sezioni di chiusura considerate all'interno della regione costituita dai bacini dei fiumi Enza, Secchia, Panaro e Reno. ....	68
Tabella II- 3: Elenco dei parametri della versione del modello HBV che si adotta. ( <b><i>TR</i></b> e <b><i>TS</i></b> non vengono calibrati).....	76
Tabella II- 4: Elenco dei bacini donatori ordinati secondo distanza di Ghosh crescente dal bacino target. Nella cross-calibrazione effettuata su bacini non annidati si studiano unicamente i donatori identificati come B.N.A.....	96
Tabella II- 5: Regione "Emilia", elenco dei bacini donatori ordinati secondo distanza di Ghosh crescente dal bacino target. B.N.A indica i bacini considerati per le analisi sui soli bacini non annidati. .	99

I'd like to thank everyone I met on my path and particularly  
thanks to my advisor, prof. Alberto Montanari,  
thanks to prof. Günter Blöschl and to Juraj Parajka  
for the precious advice!

Many many thanks to prof. Jasper Vrugt as well, who  
kindly provided us with the DREAM code.